

2019

Enabling Robust State Estimation through Covariance Adaptation

Ryan Magnum Watson
West Virginia University, rwatso12@gmail.com

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>



Part of the [Navigation, Guidance, Control and Dynamics Commons](#)

Recommended Citation

Watson, Ryan Magnum, "Enabling Robust State Estimation through Covariance Adaptation" (2019).
Graduate Theses, Dissertations, and Problem Reports. 7455.
<https://researchrepository.wvu.edu/etd/7455>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Enabling Robust State Estimation through Covariance Adaptation

RYAN M. WATSON

DISSERTATION SUBMITTED TO THE
BENJAMIN M. STATLER COLLEGE OF ENGINEERING AND MINERAL RESOURCES
AT WEST VIRGINIA UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN
AEROSPACE ENGINEERING

JASON N. GROSS, PH.D., CHAIR
YU GU, PH.D.
ROBERT C. LEISHMAN, PH.D.
DAVID S. MEBANE, PH.D.
NATALIA A. SCHMID, PH.D.
CLARK N. TAYLOR, PH.D.

DEPARTMENT OF MECHANICAL AND AEROSPACE ENGINEERING

MORGANTOWN, WEST VIRGINIA

2019

KEYWORDS: ROBUST STATE ESTIMATION, COVARIANCE ADAPTATION, GNSS

COPYRIGHT © 2019 - RYAN M. WATSON

CC BY-NC 4.0

ABSTRACT

Enabling Robust State Estimation through Covariance Adaptation

Ryan M. Watson

Several robust state estimation frameworks have been proposed over the previous decades. Underpinning all of these robust frameworks is one dubious assumption. Specifically, the assumption that an accurate *a priori* measurement uncertainty model can be provided. As systems become more autonomous, this assumption becomes less valid (i.e., as systems start operating in novel environments, there is no guarantee that the assumed *a priori* measurement uncertainty model characterizes the sensors current observation uncertainty).

In an attempt to relax this assumption, a novel robust state estimation framework is proposed. The proposed framework enables robust state estimation through the iterative adaptation of the measurement uncertainty model. The adaptation of the measurement uncertainty model is granted through non-parametric clustering of the estimator's residuals, which enables the characterization of the measurement uncertainty via a Gaussian mixture model. This Gaussian mixture model based measurement uncertainty characterization can be incorporated into any non-linear least square optimization routine.

Within this dissertation, the proposed framework is instantiated into three novel robust state estimation algorithms: batch covariance estimation (BCE), batch covariance estimation over an augmented data space (BCE-AD), and incremental covariance estimation (ICE). To verify the proposed framework, three global navigation satellite system (GNSS) data sets were collected. The collected data sets provide varying levels of observation degradation to enable the characterization of the proposed algorithm on a diverse data set. Utilizing these data sets, it is shown that the proposed framework exhibits improved state estimation accuracy when compared to other robust estimation techniques when confronted with degraded data quality.

Contents

1	INTRODUCTION	1
1.1	Problem Description	2
1.2	Dissertation Contributions	4
1.3	Dissertation Outline	5
2	LITERATURE REVIEW	8
2.1	State Estimation	9
2.1.1	Traditional State Estimation	9
2.1.2	The Limitations	12
2.1.3	Robust State Estimation	13
2.1.4	The Limitations	14
3	STATE ESTIMATION	15
3.1	Batch State Estimation	16
3.2	Incremental State Estimation	19
3.3	Robust State Estimation	20
3.3.1	M-Estimators	21
3.3.2	Switch Constraints	24
3.3.3	Dynamic Covariance Scaling	25
3.3.4	Max Mixtures	26
3.4	Key Takeaways	27
4	GPS SENSOR MODEL	29
4.1	Chapter Abstract	30
4.2	Literature Review	31
4.3	GPS Based State Estimation	33
4.3.1	Constructing the Factor Graph Constraints	33
4.3.2	Incremental Factor Graph Inference	39
4.4	Experimental Evaluation	43

4.4.1	Evaluation on Simulated Data	43
4.4.2	Evaluation on Collected Data	47
4.5	Key Takeaway	50
5	BATCH COVARIANCE ESTIMATION (BCE)	52
5.1	Chapter Abstract	53
5.2	BCE Algorithm Formulation	53
5.2.1	Data Model	54
5.2.2	Gaussian Mixture Model Fitting	58
5.2.3	Algorithm Overview	60
5.3	Experimental Validation	63
5.3.1	Data Collection	63
5.3.2	GPS Factor Graph Model	67
5.3.3	Results	69
5.3.4	<i>A Priori</i> Information Sensitivity Comparison	75
5.3.5	Run-Time Comparison	77
5.4	Key Takeaway	78
6	BCE IN AN AUGMENTED DATA SPACE (BCE-AD)	80
6.1	Chapter Abstract	81
6.2	BCE-AD Algorithm Formulation	82
6.2.1	The Augmented Data Space	82
6.2.2	The Data Model	82
6.2.3	Variational Clustering	84
6.2.4	Feature Selection	85
6.2.5	Algorithm Overview	86
6.3	Experimental Validation	89
6.3.1	Data Collection	89
6.3.2	Results	90
6.4	Key Takeaway	95
7	INCREMENTAL COVARIANCE ESTIMATION (ICE)	96
7.1	Chapter Abstract	97
7.2	The ICE Algorithm Formulation	97
7.2.1	Data Model	98
7.2.2	Uncertainty Model Adaptation	99
7.2.3	Algorithm Overview	103
7.3	Results	105
7.3.1	Data Collection	105

7.3.2	Evaluation	106
7.4	Key Takeaway	112
8	CONCLUDING REMARKS	114
8.1	Discussion	115
8.2	Contributions	116
8.3	Future Work	118
8.3.1	How can we most accurately estimate the uncertainty model?	118
8.3.2	How general are the learned uncertainty models?	118
8.3.3	What metadata is useful in other domains?	119
	REFERENCES	129

List of Figures

3.3.1	The cost function response for two commonly utilized robust maximum likelihood estimators. Robust in this sense implies that the cost functions increases more slowly than the l^2 -norm for large values of x . The specific form of the robust maximum likelihood estimators is presented in Table 3.3.1.	22
4.3.1	GNSS based localization error as a function of the utilized ionosphere model. . .	34
4.3.2	GNSS based localization error as a function of the utilized troposphere model. . .	35
4.3.3	GNSS based localization error as a function of the utilized satellite orbit and clock model.	36
4.3.4	Example sparse measurement Jacobian for the PPP processing strategy. Figure (A) shows the measurement Jacobian when a new carrier-phase ambiguity state is added for each epoch, where N.Z. represents the number of non-zero elements. Figure (B) shows the measurement Jacobian when a new carrier-phase ambiguity is added only when a new satellite is tracked or if a carrier-phase cycle-slip occurs, where N.Z. is the number of non-zero elements.	39
4.3.5	GNSS factor graph construction where the carrier-phase ambiguity states are treated as “land-mark” observations.	40
4.3.6	Generating a Bayes Net from the original factor graph, as presented within Fig. 4.3.5, using the elimination ordering, $\{B_1, B_2, X_1, X_2, X_3\}$	41
4.3.7	Generating the Bayes tree from cliques in the chordal Bayes net, as constructed within Fig. 4.3.6.	42
4.4.1	Example flight trajectory utilized for simulated GNSS observation generation. . .	44
4.4.2	Example RSOS positioning error profile for a typical simulated data set	45
4.4.3	Example phase bias convergence rate for a typical simulated data set	46
4.4.4	CDF of the RSOS positioning error for all epochs over the 100 simulated datasets.	47
4.4.5	CDF of the RSOS positioning error for epochs within the convergence period over the 100 simulated datasets.	48
4.4.6	<i>Phastball</i> research platform [1] in flight over the West Virginia University Jackson’s Mill airfield.	48

4.4.7	Flight profile for collected data set.	49
4.4.8	RSOS positioning error comparison for a Kalman filter and the incremental graph optimizer when a Phastball collected dataset is utilized.	50
5.2.1	Graphical representation of the utilized data model which visually encodes the joint probability distribution presented in Eq. 5.6. Each red node represents a latent parameter, and the blue node represents the provided observations. The plate notation [2] is utilized to represent the replication of a random variable.	57
5.2.2	Schematical overview of the batch covariance estimation algorithm. The proposed approach enables robust estimation through the iterative calculation of the measurement error covariance model, where the measurement error covariance model is characterized by a Gaussian mixture model that is fit to the state estimation residuals.	62
5.3.1	Ground trace of three kinematic driving data sets collected in the Morgantown, WV area.	63
5.3.2	Box plot of RSOS positioning error for collected kinematic GNSS data sets with low quality – see Table 5.3.1 for GNSS receiver configuration – generated observations. The specific estimator statistics are provided in Table 5.3.3. Within this figure, L_2 , is a batch estimator with l^2 cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and BCE is the proposed batch covariance estimation technique.	71
5.3.3	Evolution of the GNSS observation Gaussian mixture model based uncertainty model for data collect 1 with low-quality observation as a function of optimizer iteration.	72
5.3.4	Progression of GNSS observation residuals with multimodal measurement error covariance – as estimated by the proposed approach which utilizes variational inference to generate a Gaussian mixture model uncertainty representation from the estimators residuals – for data collect 1 with low-quality observations.	72
5.3.5	Box plot of RSOS positioning error for collected kinematic GNSS data sets with high quality – see Table 5.3.1 for GNSS receiver configuration – generated observations. The specific estimator statistics are provided in Table 5.3.4. Within this figure, L_2 , is a batch estimator with L_2 cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and BCE is the proposed batch covariance estimation technique.	74

5.3.6	Measurement residuals with associated BCE estimated measurement error covariance – each ellipse represents 95% confidence – for data collect 1 with high quality observations.	75
5.3.7	Sensitivity of the estimation frameworks to the selection of an <i>a priori</i> measurement error covariance model. Within this figure, L_2 is a batch estimator with l^2 cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and BCE is the proposed batch covariance estimation technique.	76
5.3.8	Wall-clock time per utilized observation for each of the estimation frameworks, where L_2 is a batch estimator with l^2 cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and BCE is the proposed batch covariance estimation technique.	78
6.2.1	Ground trace for the three kinematic GNSS data collects. All three data collected were recorded in Morgantown, Wv	84
6.2.2	Overview of the proposed robust optimized algorithm. The proposed approach enables robust state estimation through the iterative estimation of the – possible multimodal – measurement error covariance model, where the measurement error covariance model is estimated by clustering over an augmented data space constructed from the previous iteration of optimization.	88
6.3.1	Ground trace for the three kinematic GNSS data collects. All three data collected were recorded in Morgantown, Wv	89
6.3.2	Horizontal RSOS positioning error for collected GNSS data sets. The specific estimator statistics are provided in Table 6.3.1. Within this figure, L_2 , is a batch estimator with L_2 cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, BCE is the batch covariance estimation technique, and BCE-AD is the proposed batch covariance estimation technique over an augmented data space.	91
6.3.3	Partitioning of the measurement residual domain with the BCE approach (see Figs. 6.3.3a, 6.3.3c, and 6.3.3e), and the BCE-AD approach (see Figs. 6.3.3b, 6.3.3d, and 6.3.3f), for the three kinematic GNSS data collects at the final iteration of optimization.	94
7.2.1	Graphical depiction of the proposed incremental covariance estimation (ICE) algorithm. The proposed approach enables efficient, incremental, and robust state estimation through the iterative adaptation of the measurement uncertainty model, based upon the state estimation residuals that correspond to erroneous observations.	104

7.3.1	Ground trace for the three utilized GNSS data sets. The white trace corresponds to data collect 1, the green trace corresponds to data collect 2, and the blue trace corresponds to data collect 3.	106
7.3.2	Incrementally estimated measurement error covariance model when the observations are generated with high fidelity receiver tracking parameters.	110
7.3.3	Estimator update time for each of the estimation frameworks over all data collects, where L_2 is a batch estimator with l^2 -norm cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and ICE is the proposed incremental covariance estimation technique.	112

List of Tables

3.3.1	Functions utilized to implement the l^2 -norm, Cauchy, and Huber cost functions, where, where $\rho(x)$ is the robust cost function, $\psi(x) \triangleq \frac{\partial \rho(x)}{\partial x}$ is the influence function, $w(x) \triangleq \frac{\psi(x)}{x}$ is the weighting function, and k is the user defined kernel width.	22
3.3.2	Equivalent dynamic covariance scaling m-estimator, $w(x)$ is the weighting function – as specified in Eq. 3.18, $\psi(X) \triangleq x w(x)$ is the influence function, $\rho(X) = \int \psi(x) dx$ is the robust cost function, and k is the user defined kernel width.	26
4.4.1	Description of randomly initialized parameters within the Monte-Carlo evaluation.	44
4.4.2	Stochastic model parameters for estimators.	45
4.4.3	Positioning statistics for a single flight	45
4.4.4	Positioning statistics for all epochs.	47
4.4.5	Positioning statistics for epochs within the convergence period.	48
4.4.6	RSOS localization statistics for collected dataset.	50
5.3.1	GPS SDR observation tracking configurations utilized within this study.	66
5.3.2	Zero-baseline observation comparison for kinematic data sets – see Fig. 7.3.1 – with varied GPS SDR tracking configurations, as specified in Table 5.3.1.	66
5.3.3	Horizontal RSOS positioning error results when low quality observations are utilized. The green and red cell entries correspond to the minimum and maximum statistic, respectively.	71
5.3.4	Horizontal RSOS positioning error results when high quality observations are utilized. The green and red cell entries correspond to the minimum and maximum statistic, respectively.	74
6.3.1	Horizontal RSOS positioning error results when low quality observations are utilized. The green and red cell entries correspond to the minimum and maximum statistic, respectively.	91

6.3.2	Utilized features for each iteration of optimization with the BCE-AD approach. The green and red cell entries correspond to the utilized and non-utilized features, respectively. Within the table, ρ is the pseudorange residual, Φ is the carrier-phase residual, El. is the elevation angle, Az. is the azimuth angle, and SS. is the signal strength.	93
7.3.1	Horizontal RSOS positioning error results when low fidelity receiver tracking parameters are utilized to generate the observations. The green and red cell entries correspond to the minimum and maximum statistic, respectively.	108
7.3.2	Horizontal RSOS positioning error results when high fidelity receiver tracking parameters are utilized to generate the observations. The green and red cell entries correspond to the minimum and maximum statistic, respectively.	109

Contributing Papers

A portion of the results and discussion presented within this document were originally contained within other articles. The articles that contribute to specific chapters are presented below.

- Ch. 1** **Watson, R.M.** , Gross, J.N., Taylor, C.N., and Leishman, R.C., 2019, September. **Enabling Robust State Estimation through Measurement Error Covariance Adaptation.** Accepted to IEEE Transactions on Aerospace Electronic Systems, September 2019.
- Ch. 2** **Watson, R.M.** and Gross, J.N., 2017, September. **Robust Navigation In GNSS Degraded Environment Using Graph Optimization.** Proceedings of the 30th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2017), Portland, Oregon, September 2017, pp. 2906-2918
- Watson, R.M.** and Gross, J.N., 2018, April. **Evaluation of kinematic precise point positioning convergence with an incremental graph optimizer.** In 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS) (pp. 589-596). IEEE.
- Watson, R.M.** , Gross, J.N., Taylor, C.N., and Leishman, R.C., 2019, September. **Enabling Robust State Estimation through Measurement Error Covariance Adaptation.** Accepted to IEEE Transactions on Aerospace Electronic Systems, September 2019.
- Ch. 3** **Watson, R.M.** , Taylor, C.N., Leishman, R.C., and Gross, J.N., 2019, October. **Robust Incremental State Estimation through Covariance Adaptation.** Submitted to IEEE Robotics and Automation Letters, October 2019.
- Watson, R.M.** , Gross, J.N., Taylor, C.N., and Leishman, R.C., 2019, September. **Enabling Robust State Estimation through Measurement Error Covariance Adaptation.** Accepted to IEEE Transactions on Aerospace Electronic Systems, September 2019.
- Ch. 4** **Watson, R.M.** and Gross, J.N., 2018, April. **Evaluation of kinematic precise point positioning convergence with an incremental graph optimizer.** In 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS) (pp. 589-596). IEEE.

- Ch. 5** **Watson, R.M.** , Taylor, C.N., Leishman, R.C., and Gross, J.N., 2018, September. **Batch Measurement Error Covariance Estimation for Robust Localization.** In Proceedings of the 31st International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2018), Miami, Florida, September 2018, (pp. 2429-2439).
- Watson, R.M.** , Gross, J.N., Taylor, C.N., and Leishman, R.C., 2019, September. **Enabling Robust State Estimation through Measurement Error Covariance Adaptation.** Accepted to IEEE Transactions on Aerospace Electronic Systems, September 2019.
- Ch. 6** **Watson, R.M.** , Gross, J.N., Taylor, C.N., and Leishman, R.C., 2019, July. **Uncertainty Model Estimation in an Augmented Data Space for Robust State Estimation.** Submitted to IEEE Transactions on Aerospace Electronic Systems, July 2019.
- Ch. 7** **Watson, R.M.** , Gross, J.N., Taylor, C.N., and Leishman, R.C., 2019, October. **Robust Incremental State Estimation through Covariance Adaptation.** Submitted to IEEE Robotics and Automation Letters, October 2019.

Contributing Software

To generate the results presented within this dissertation, several software packages have been developed. The developed software has been made publicly available in the hope that it will be found beneficial in enabling future research. The specific packages utilized for a given chapter is provided below.

- Ch. 4**
 - **Package Name:** PPP-BayesTree
 - **Package Link:** <https://github.com/wvu-navLab/PPP-BayesTree>
 - **Data Link:**
<https://github.com/wvu-navLab/PPP-BayesTree/tree/master/trunk/data>

- Ch. 5**
 - **Package Name:** BCE
 - **Package Link:** <https://github.com/wvu-navLab/Enabling-Robust-State-Estimation-through-Measurement-Error-Covariance-Adaptation>
 - **Data Link:** <https://github.com/wvu-navLab/GnssData>

- Ch. 6**
 - **Package Name:** BCE-AD
 - **Package Link:** https://github.com/wvu-navLab/BCE_AD
 - **Data Link:** <https://github.com/wvu-navLab/GnssData>

- Ch. 7**
 - **Package Name:** ICE
 - **Package Link:** <https://github.com/wvu-navLab/ICE>
 - **Data Link:** <https://github.com/wvu-navLab/GnssData>

List of Acronyms

GPS	global positioning system	63
GNSS	global navigation satellite system	3
PLL	phase lock loop	105
DLL	delay lock loop	64
IQ	in-phase and quadrature	63
SDR	software defined radio	63
RINEX	receiver independent exchange format	66
RTK	real time kinematic	89
PPP	precise point positioning	31
EKF	extended Kalman filter	11
iSAM₂	incremental smoothing and mapping	20
GMM	Gaussian mixture model	4
LS	least squares	17
NLLS	nonlinear least squares	2
SLAM	simultaneous localization and mapping	6
SAM	smoothing and mapping	4
LARS	least angel regression	86
m-estimator	maximum likelihood type estimator	3
IRLS	iteratively re-weighted least squares	23
DCS	dynamic covariance scaling	3
MM	max-mixtures	13

BCE batch covariance estimation	60
ICE incremental covariance estimation	97
BCE-AD batch covariance estimation over an augmented data-space	82
RANSAC random sample consensus	3
RRR realizing, reversing, recovering	3
RAIM receiver autonomous integrity monitoring	3
RSOS residual-sum-of-squares	90
NN nearest neighbor	86

Acknowledgments

To enable me to get to this point, I have been lucky to receive assistance from many great friends and mentors. Thus, to show my appreciation, I would like to take a brief moment to acknowledge a few of them.

First, I would like to thank **Dr. Jason Gross** for being a phenomenal advisor over my years as a graduate student. This work certainly would not have been possible without your mentorship. I would also like to thank **Dr. Yu Gu** particularly for piquing my interest in robotics as an undergraduate student.

The inception of the majority of the ideas presented within this document were formed while conducting an internship with **Dr. Clark Taylor** and **Dr. Robert Leishman**. Working with both of you over the course of my graduate education has been highly influential. I greatly appreciate the time and effort that both of you have put into this research.

Additionally, I would like to thank the remaining members of my committee, **Dr. David Mebane**, and **Dr. Natalia Schmid**. You both have been incredible teachers, and thus greatly shaped me as a researcher.

Finally, I would like to thank the **Autonomy and Navigation Technology (ANT) Center** at the Air Force Institute of Technology for funding the research presented within this document.

1

Introduction

1.1 PROBLEM DESCRIPTION

The applicability of robotic platforms to an ever increasing number of applications (e.g., disaster recovery [3], scientific investigations [4], health care [5]) has been realized in recent years. A core component that enables the operation of these platforms is the ability to localize (i.e., estimate the position states within a given coordinate frame) given *a priori* information and a set of measurements. Thus, a considerable amount of research has been afforded to the problem of accurate and robust localization of a robotic platform.

To facilitate the localization of a platform, a state estimation [6] scheme must be implemented. These state estimation frameworks can take two forms: either a batch estimator, or an incremental estimator. The batch variant of state estimation is generally facilitated through the utilization of a nonlinear least squares (NLLS) algorithm (e.g., line-search method [7], or a trust-region method [8]). On the other hand, incremental state estimators are usually implemented as a variant of the Kalman filter [9] (e.g., the extended Kalman filter [10], or the unscented Kalman filter [11]).

The cost function underlying all of the state estimation techniques mentioned above is the l^2 -norm of the estimation errors. This cost function enables accurate and efficient estimation when the assumed models accurately characterize the observations. However, when the utilized model does not accurately characterize the provided measurements, the estimation framework can provide an arbitrarily poor state estimate. Specifically, as discussed in [12], the l^2 -norm cost function has an asymptotic breakdown of zero (i.e., if any single observation deviates from the utilized model, the estimated solution can be biased by an arbitrarily large quantity [13]).

To combat the poor breakdown properties of the l^2 -norm cost function, several robust estimation frameworks have been proposed. These frameworks can be broadly partitioned into two cat-

egories: data weighting methods, and data exclusion methods. The data weighting methods work by iteratively calculating a measurement weighting vector such that observations which most substantially deviated from the utilized model have a reduced influence on the estimated result. Several commonly utilized data weight techniques include: robust maximum likelihood type estimators (m-estimators) [14], switchable constraints [15], dynamic covariance scaling (DCS) [16], and max mixtures [17]. On the other hand, the data exclusion methods work by finding a trusted subset of the provided observations. Several commonly utilized data exclusion techniques include: random sample consensus (RANSAC) [18], realizing, reversing, recovering (RRR) [19], l^1 relaxation [20], and receiver autonomous integrity monitoring (RAIM) (e.g., solution separation [21] or residual-based [22] techniques¹).

Both estimation paradigms have certain undesirable properties. For example, in many applications, it is undesirable to completely remove observations that do not adhere to the defined model (e.g., the utilization of global navigation satellite system (GNSS) in an urban environment where it may not be possible to estimate the desired states if observations are removed). Instead, it can be more informative to accurately characterize the measurement uncertainty model and accordingly reduce the observation's influence on the estimate. This concept of data uncertainty model characterization naturally conforms to the data weighting class of robust techniques (as discussed more thoroughly in Chapter 3). However, the currently utilized weighting techniques have the undesirable property that they assume an accurate characterization of the true measurement uncertainty model can be provided *a priori*, which may not be a valid assumption if the robotic platform is operating in an environment that has not been previously experienced.

Within this dissertation, a novel robust data weighting methodology is developed, which relaxes

¹For a thorough comparison between the two commonly utilized RAIM implementations, the reader is referred to [23].

the assumption that an accurate *a priori* measurement uncertainty characterization can be provided. This assumption is relaxed by iteratively estimating a Gaussian mixture model (GMM) based on the state estimation residuals that characterizes the measurement uncertainty model. The approach provided in this dissertation is evaluated on several collected GNSS data sets. Utilizing these data sets, it is shown that the proposed methodology exhibits improved state estimation accuracy when compared to other robust estimation techniques when confronted with degraded data quality.

1.2 DISSERTATION CONTRIBUTIONS

The work detailed within this dissertation, can be summarized through three primary contributions;

Contribution 1: A novel GNSS signal processing technique is developed. Specifically, a GNSS signal processing strategy that exploits the smoothing and mapping (SAM) framework commonly utilized within the robotic community is developed. The developed strategy leverages a connection made, within this dissertation, between the carrier-phase ambiguity states for GNSS signal processing and the loop-closure constraints which are common to the SAM formulation.

Contribution 2: A novel robust state estimation framework is developed. Specifically, a novel robust state estimation framework that works through the iterative adaptation of the measurement uncertainty model is developed. The adaptation of the measurement uncertainty model is achieved through non-parametric clustering, which enables the characterization of the measurement uncertainty via a Gaussian mixture model. The provided Gaussian mixture model can be utilized within any non-linear least squares optimization algorithm by

approximately characterizing each observation with the sufficient statistics of the assigned cluster (i.e., each observation's uncertainty model is updated based upon the assignment provided by the non-parametric clustering algorithm). Within this dissertation, both a batch (i.e., the batch covariance estimation (BCE), as detailed in Chapter 5) and incremental (i.e., the incremental covariance estimation (ICE), as detailed in Chapter 7) version of the proposed methodology is developed.

Contribution 3: A novel methodology for the incorporation of metadata into the state-estimation framework is developed. Specifically, it is argued that the uncertainty estimation process should be augmented to include metadata (e.g., the signal strength of the associated GNSS observation). The modification of the uncertainty estimation process to an augmented data space is significant because it increases the likelihood of a unique partitioning in the measurement residual domain and thus provides the ability to more accurately characterize the measurement uncertainty model.

1.3 DISSERTATION OUTLINE

To enable a more thorough discussion of the contributions discussed in the previous section (i.e., Section 1.2), the remainder of this dissertation is organized in the following manner.

Chapter 2: Within this chapter, a literature review for the subject discussed within this dissertation is provided. This review begins with an overview traditional state estimation and its robust variants. Then, the focus of the review turns to an overview of GMM fitting techniques.

Chapter 3: Within this chapter, a succinct overview of state estimation is provided. This

review begins with an overview of traditional state estimation (i.e., non-robust estimation). Then, a review of robust state estimation methodologies, with the primary objective of the review being to provide a unifying view of robust state estimation methodologies as enabling robustness by adapting the *a priori* covariance model is provided.

Chapter 4: Within this chapter, a thorough review of the GNSS sensor model which will be utilized throughout this dissertation is provided. This review will begin with a succinct review of the GNSS observations, and their primary error sources. Then, a novel GNSS estimation formulation which details the parallels between GNSS signal processing and simultaneous localization and mapping (SLAM) optimization with loop-closure constraints is detailed.

Chapter 5: Within this chapter, a batch covariance estimation technique, which enables robust state estimation through the iterative adaptation of the measurement uncertainty model is developed. The proposed algorithm is verified on three GNSS collected data sets, where it is shown that the proposed technique exhibits some advantages when compared to other robust estimation techniques when confronted with degraded data quality.

Chapter 6: Within this chapter, a batch covariance estimation technique over an augmented data space is presented. This approach extends the work proposed in Chapter 5 by arguing that the uncertainty estimation process should be augmented to include metadata (e.g., the signal strength of the associated GNSS observation). The proposed algorithm is evaluated against the several state-of-the-art robust estimators, where it is shown that a significant increase in state estimation accuracy can be granted.

Chapter 7: Within this chapter, an incremental extension of the previously developed batch

covariance estimation technique is presented. Specifically, an efficient, incremental extension of the methodology is proposed. The incremental covariance estimation approach, as detailed within this chapter, is evaluated on three collected data sets, where it is shown to provide a significant increase in localization accuracy, when compared to other robust estimators.

Chapter 8: Within this chapter, the dissertation concludes. This chapter begins with an overview of the topics covered within this work, and concludes with proposed future research directions.

2

Literature Review

The ability to precisely localize a platform is of paramount importance to a myriad of fields (e.g., augmented reality [24], autonomous navigation [25], and natural hazard monitoring [26]). As these systems become fully autonomous (i.e., without human intervention), there is a need to develop robust localization algorithms. As a step toward enabling this, as discussed within this dissertation, a novel robust estimation framework is developed.

To enable a discussion of the proposed robust estimation framework, first, the foundation is laid for the framework through a literature review. This review will begin with a discussion on traditional state estimation. Then, the discussion turns to a brief overview of the current robust variants of traditional state estimation.

2.1 STATE ESTIMATION

To begin the literature review, a succinct review of state estimation is provided. This review will commence with an overview of the advances made within the traditional (i.e., based upon the l^2 -norm cost function) state estimation community. This review of traditional state estimation will cover both the filtering paradigm [6], and the SAM paradigm [27]. Next, the review will turn to the field of robust state estimation where a concise overview of the recent advances within the field, with specific emphasis placed on robust state estimation contributions commonly adopted within the robotics community, is detailed.

2.1.1 TRADITIONAL STATE ESTIMATION

Traditional state estimation is concerned with estimating a set of states (i.e., a set of parameters which define the system of interest) through the minimization of a cost function that is entirely constructed via the l^2 -norm. Techniques to enable this type of estimation can broadly be parti-

tioned based upon the amount of information they marginalize at each time-step. Specifically, there is the filtering paradigm, which generally marginalizes out all prior information after each update. And, there is the smoothing paradigm, which retains prior information (i.e., does not marginalize out all prior information), to allow for re-linearization of prior state estimates as new information is obtained.

THE FILTERING PARADIGM

The filtering paradigm builds upon a rich history of state estimation that dates to Karl Friedrich Gauss. Specifically, the filtering paradigm builds upon the concept of finding an optimal set of states through the minimization of the mean squared error, as introduced by Gauss. This process of minimizing the mean squared error in a batch setting is generally termed least-squares, and has been thoroughly studied since its inception. For a review of this incremental development, the reader is referred to [28].

In modern times (i.e., since its inception in the early 1960's), problems that fall within the domain of filtering are generally addressed through the application of a variant of the traditional Kalman filter [9]. The Kalman filter works by finding an optimal – where optimality is quantified by minimizing the mean squared error [6] – combination of the predicted and observed values in an incremental nature.

However, this framework is only valid under a set of restrictive assumptions. The primary assumption that restricts the applicability of the traditional Kalman filter is the assumption that the system adheres to a linear model; however, additional assumptions are also enforced, such as, the noise on the observations and the system dynamics are independent. For a review of the theory and implementation of the Kalman filter, the reader is referred to [29, 30].

To extend this work to the non-linear regime, Smith introduced the extended Kalman filter (EKF) [31]. The EKF works by linearizing the non-linear system about its nominal trajectory (i.e., taking a first order Taylor series expansion of the systems about the trajectory [6]). It should be noted that other non-linear filter techniques have been proposed (i.e., the unscented Kalman filter [32], or the particle filter [33]); however, in practice, the EKF seems to be the most widely utilized filter technique, with applications spanning domains from aerospace navigation [34] to ecological modelling [35].

THE SMOOTHING AND MAPPING PARADIGM

While the filtering paradigm works well for many applications, the marginalization of all prior information is not advantageous to all applications (e.g., SLAM where re-linearization of prior states is desired when new loop-closure information is available [36]). To enable the re-linearization of prior states when new information is present, the smoothing [37] framework is utilized. Since the seminal paper on the subject, [38], research on smoothing has been dominated by probabilistic graphical methodologies [39].

When [38] was published, graph-based smoothing was not widely utilized due to computation complexity of solving the initial formulation. However, quickly thereafter, methods were proposed to greatly reduce complexity through the utilization of factor graphs [27, 40]. The $\sqrt{\text{SAM}}$ formulation as presented in [41] was particularly influential as it drew connections between the factor graph formulation and sparse linear algebra (i.e., they were able to show that the factor graph directly encodes the measurement Jacobian matrix [41] and thus, can efficiently be solved via a NLLS algorithm [27].).

This factor graph formulation of smoothing was later extended to enable incremental updates

(i.e., incremental SAM (iSAM)) in [42]. The iSAM formulation enables incremental updates through appending the previously factorized system (i.e., the QR-factorization [43] of the information matrix) when new information is present. Then, Givens rotations were utilized to ensure that the factorization maintained its structure (i.e., an upper-triangular system).

The iSAM formulation required batch re-linearization of the entire matrix factorization to ensure sparsity [42]. To mitigate this batch re-computation, the Bayes tree [44] was introduced. This directed graphical model directly represents the square root information matrix and can be easily computed from the associated factor graph in a two-step process, as detailed in [45]. Due to the structure of the Bayes tree graphical model, this methodology removes the requirement to re-factor the entire system when new information is added. Instead, only the affected section of the Bayes tree is re-factored, as detailed within [45].

2.1.2 THE LIMITATIONS

One major disadvantages of a traditional state estimator – indifferent to the utilized framework (i.e., a filter of a SAM methodology) – is the undesirable asymptotic breakdown property [13] as a result of utilizing a l^2 -norm cost function exclusively. Any estimator that exclusively utilizes the l^2 -norm to construct the cost function will have an asymptotic breakdown point of zero. As discussed in [12], this property can be intuitively understood by letting any arbitrary observation, y_n , significantly deviate from the model (i.e., $\|y_n - h_n(X)\|_{\Lambda_n} \rightarrow \infty$), which, in turn, will arbitrarily bias the provided state estimate (i.e., $\|X\| \rightarrow \infty$). To mitigate this issue, the discuss will next progress to an overview of several robust estimation techniques.

2.1.3 ROBUST STATE ESTIMATION

As discussed, the traditional state estimation methodologies provide optimal results when the collected observations adhere to the *a priori* models. However, in many robotic applications of interest, the observations can be degraded (e.g., GNSS observations in an urban environment, or RGB observations in a low-light setting), which cause a deviation between the collected observations and the assumed models. When this deviation is present, the traditional state estimation schemes (i.e., estimators that utilize the l^2 -norm exclusively to construct the cost-function) can breakdown [13].

To overcome the breakdown of traditional state estimators in data degraded scenarios, several robust estimation schemes have been developed. These robust estimation schemes reduce the effect that erroneous observations have on the estimation process by scaling the associated covariance matrix [46]. To enable this covariance scaling in practice, several implementations have been developed (i.e., m-estimators [14], switchable constraints [47], and DCS [16]).

To extend robust state estimation from the traditional uni-modal uncertainty model paradigm to a multi-modal implementation, the max-mixtures (MM) [17] approach was developed. The MM approach mitigates increased computation complexity generally assumed to accompany the incorporation of multi-modal uncertainty models by first assuming that the uncertainty model can be represented by a GMM, then selecting the single Gaussian component from the GMM that maximizes the likelihood of the individual observation given the current state estimate.

For a more thorough discussion on implementation of the above discussed robust estimators, the reader is referred to Chapter 3 of this dissertation. Within this chapter, a discussion of the above mentioned robust estimators, from the unifying view of enabling robust state estimation through covariance adaptation, is provided.

2.1.4 THE LIMITATIONS

As discussed, several robust state estimation frameworks have been proposed. Each of the discussed estimators work under the assumption that an accurate *a priori* measurement error covariance is available. However, in practice, the requirement to supply an accurate *a priori* characterization of the measurement uncertainty model is not always feasible when considering that the platform could be operating in a novel environment, a non-cooperative environment, or both. To relax this assumption, a novel robust estimation framework is developed within this dissertation that attempts to learn the covariance model in conjunction with the desired states. The novel framework is proposed within Chapter 5, and extended within Chapters 6 and 7.

3

State Estimation

Within this chapter, an understanding of the traditional state estimation paradigm will be developed. Specifically, this understanding will be developed from the SAM point-of-view, as it is what will be utilized through this dissertation. To foster this understanding, this Chapter will begin an overview of the batch estimation framework, as detailed in Section 3.1. Then, an overview of incremental state estimation is provided in Section 3.2. Finally, a review of robust state estimation is provided in Section 3.3.

3.1 BATCH STATE ESTIMATION

The problem generally termed state estimation is primarily concerned with finding the set of states X (i.e., a set of parameters that describe the system of interest) that is in accordance with the set of provided information Y . To evaluate the level of accordance between the set of states and the provided information, it is common to utilize the conditional distribution presented in Eq. 3.1 (i.e., the optimal state estimate \hat{X} is the state vector that maximizes the probability of the set of states conditioned on the provided information).

$$\hat{X} = \underset{x}{\operatorname{argmax}} p(X | Y) \quad (3.1)$$

To enable the efficient representation of this estimation problem, the factor graph [27] has been extensively utilized. This representation is utilized because it enables the factorization of the complex *a posteriori* distribution into the product of simplified functions, as presented in Eq. 3.2.

Where, within Eq. 3.2, ψ_n is a factor, $A_n \subseteq \{X_1, X_2, \dots, X_N\}$, and $B_n \subseteq \{Y_1, Y_2, \dots, Y_N\}$.

$$p(X | Y) \propto \prod_{n=1}^N \psi_n(A_n, B_n), \quad (3.2)$$

With the factorization of the *a posteriori* distribution, as presented in Eq. 3.2, the state estimation problem simplifies to the canonical least squares (LS) form [27], as presented in Eq. 3.3, where h_n is the measurement function (i.e., a function that maps the state estimate to the measurement domain), Λ is the assumed uncertainty model, and $\|\cdot\|$ is the l^2 -norm. However, it should be noted that this simplification is only true if it is assumed that all of the factors within the factorization adhere to a Gaussian model.

$$\hat{X} = \underset{X}{\operatorname{argmin}} \sum_{n=1}^N \|r_n(X)\|_{\Lambda_n} \quad \text{s.t.} \quad r_n(X) \triangleq y_n - h_n(X), \quad (3.3)$$

In general, for the non-linear case, there is no direct solution to the problem presented in Eq. 3.3. Thus, an incremental methodology of the form $X_t = X_{t-1} + \hat{\Delta}_X$ must be employed. To find an incremental update to the state estimate, it is common to linearize the measurement function about the current state estimation, as presented in Eq. 3.4. The linearized representation of the estimation problem presented in Eq. 3.4 can be simplified by pulling the covariance matrix inside the norm, as presented in Eq. 3.5, where a_n and b_n are the whitened measurement Jacobian and state estimation residual vectors (i.e., $a_n \triangleq \Lambda_n^{-1/2} \frac{\partial h_n(X_{t-1})}{\partial X}$, and $b_n \triangleq \Lambda_n^{-1/2} r_n$), respectively.

$$\hat{\Delta}_X = \operatorname{argmin}_{\Delta_X} \sum_{n=1}^N \left\| \frac{\partial h_n(X_{t-1})}{\partial X} \Delta_X - r_n \right\|_{\Lambda_n} \quad (3.4)$$

$$= \operatorname{argmin}_{\Delta_X} \sum_{n=1}^N \|a_n \Delta_X - b_n\|, \quad (3.5)$$

The cost function presented in Eq. 3.5, can be more compactly defined as presented within Eq. 3.6, where the matrices A and B are defined by stacking vertically their respective whitened components (i.e., A is a matrix formed by vertically stacking the set $\{a_1, \dots, a_N\}$, and B is a matrix formed by vertically stacking the set $\{b_1, \dots, b_N\}$).

$$\hat{\Delta}_X = \operatorname{argmin}_{\Delta_X} \|A\Delta_X - B\|, \quad (3.6)$$

To solve the system presented in Eq. 3.6, it is common to utilize a matrix factorization of the measurement Jacobain matrix¹ [49]. For this discussion, the QR-decomposition [30] is utilized, which provides a factorization as presented in Eq. 3.7, where $Q \in \mathbb{R}^{N \times N}$ is an orthogonal matrix and $R \in \mathbb{R}^{M \times M}$ is an upper-triangular matrix.

$$A = Q \begin{bmatrix} R \\ \mathbf{o} \end{bmatrix}, \quad (3.7)$$

Utilizing the factorization presented in Eq. 3.7, the cost function presented in Eq. 3.6 can equiv-

¹To make a connection back to the graphical model (i.e., the factor graph), it was shown in [45] that variable elimination [48] on the factor graph (i.e., converting a factor graph to a Bayes net) is equivalent to QR-decomposition.

alently² be expressed as provided in Eq. 3.8, which simplifies to the expression provided in Eq. 3.9. The expression provided in Eq. 3.9 is computational efficient due to the upper triangular nature of the matrix R (i.e., the system $R\hat{\Delta}_X = c$ can simply be solved via back substitution).

$$\hat{\Delta}_X = \underset{\Delta_X}{\operatorname{argmin}} \|Q^T(A\Delta_X - B)\| \quad (3.8)$$

$$= \underset{\Delta_X}{\operatorname{argmin}} \|R\Delta_X - c\| + \|d\| \quad \text{s.t.} \quad Q^T B \triangleq \begin{bmatrix} c \\ d \end{bmatrix} \quad (3.9)$$

3.2 INCREMENTAL STATE ESTIMATION

The estimation framework discussed previously provides an efficient and numerically stable solution when all of the information is provided beforehand. However, for many applications, the information is provided incrementally. When this is the case, the estimation framework discussed previously is inefficient due to the need to recompute the QR-decomposition of the entire measurement Jacobian matrix every time a new information is provided.

To overcome this computation limitation, the concept of incrementally updating the QR-decomposition was studied within [50]. Within [50], incremental updating of the matrix factorization was enabled by first augmenting the previous factorization (i.e., incorporating new rows in the R and c matrices), then, restoring the upper triangular form of the factorization using Givens rotations³.

²The cost functions presented in Eq. 3.6 and Eq. 3.8 are equivalent due to the orthogonality of the matrix Q (i.e., $\|Qv\| = \|v\|$ given that Q is orthogonal).

³See section 5.1.8 of [43] for a thorough review of Givens rotations with applications to LS

The approach proposed within [50] does have one key limitation, which is the requirement to conduct periodic batch re-computation of the QR-decomposition for the entire measurement Jacobian matrix to enable variable re-ordering. This batch re-computation is utilized to maintain the sparsity of the upper-triangular system. To mitigate this batch re-computation the Bayes tree [44] was introduced. This directed graphical model directly represents the square root information matrix and can be easily computed from the associated factor graph in a two-step process, as detailed in [45]. Due to the structure of the Bayes tree graphical model, this methodology removes the requirement to re-factor the entire system when new information is added. Instead, only the affected section of the Bayes tree is re-factored, as detailed within [45]. This approach to state estimation is titled incremental smoothing and mapping (iSAM₂), and is the approach utilized within this work (e.g., see Section 4.3.2 for an application to GNSS signal processing).

3.3 ROBUST STATE ESTIMATION

One of the major disadvantages of utilizing the l^2 -norm cost function is the undesirable asymptotic breakdown property [13]. This undesirable property of the l^2 -norm cost function has spurred significant research interest in the direction of robust state estimation [14–17]. Within this Chapter, several of the major advances within the field will be discussed from the point-of-view of covariance adaptation; however, it should be noted that this is not an exhaustive compilation of all robust state estimation techniques. Rather, the purpose of this discussion is to provide a concise overview of the recent advances within the field, with specific emphasis placed on robust state estimation contributions now commonly adopted within the robotics community.

3.3.1 M-ESTIMATORS

One of the initial frameworks developed to enable robust state estimation is through a class of m-estimators, as introduced by Huber [14, 51], that are less sensitive to erroneous observations⁴ than the l^2 -norm. Specifically, this framework aims to replace the l^2 -norm cost function, with a modified cost function that is more robust (i.e., increases more slowly when compared to the l^2 -norm cost function, as depicted in Fig. 3.3.1). This class of modified cost functions, generically, take a form as presented in Eq. 3.10, where $\rho(\cdot)$ is the modified cost function.

$$\begin{aligned}\hat{X} &= \underset{X}{\operatorname{argmin}} \sum_{n=1}^N \rho(\|r_n(X)\|_{\Lambda_n}) \\ &= \underset{X}{\operatorname{argmin}} \sum_{n=1}^N \rho(e_n) \quad \text{s.t.} \quad e_n \triangleq \|r_n(X)\|_{\Lambda_n}\end{aligned}\tag{3.10}$$

While numerous robust cost functions exist, this discussion will focus on two commonly utilized functions: the Huber and Cauchy cost functions. (For additional m-estimators, the reader is referred to [52].) Associated with each robust cost function is an influence and weighting function that can be used to enable these cost functions to be solved with the same **NLLS** approaches used for a quadratic cost function. A summary of the cost, influence and weight function is shown in Table 3.3.1.

As discussed at the outset of this Chapter, our objective is to view robust state estimation as a type of covariance adaptation. To view m-estimators as a type of covariance adaptation, this exer-

⁴Erroneous observations are simply observations that do not adhere to the defined observation model.

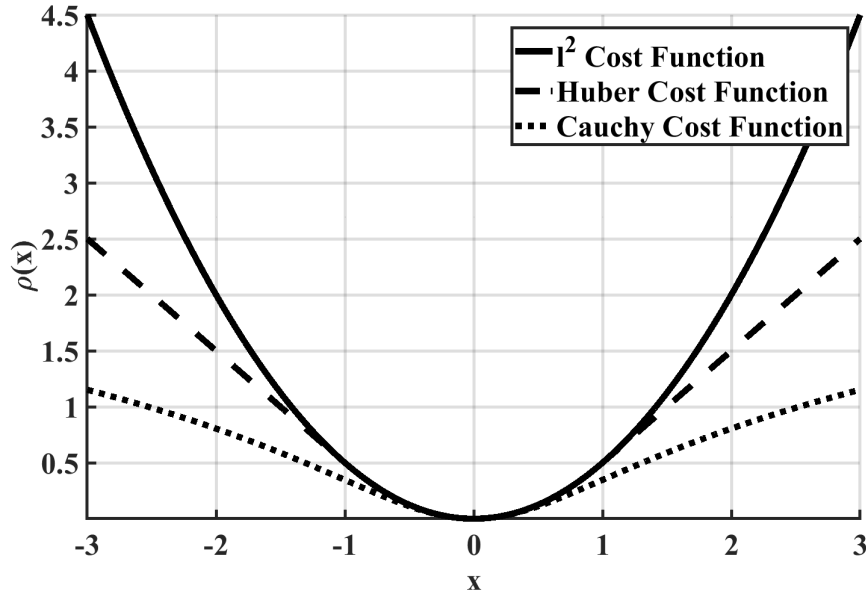


Figure 3.3.1: The cost function response for two commonly utilized robust maximum likelihood estimators. Robust in this sense implies that the cost functions increases more slowly than the l^2 -norm for large values of x . The specific form of the robust maximum likelihood estimators is presented in Table 3.3.1.

M-Estimator	$\rho(x)$	$\psi(x)$	$w(x)$
l^2 -norm	$x^2 / 2$	x	1
Huber $\begin{cases} \text{if } x \leq k \\ \text{if } x > k \end{cases}$	$\begin{cases} \frac{1}{2}x^2 \\ k(x - \frac{k}{2}) \end{cases}$	$\begin{cases} x \\ k \operatorname{sgn}(x) \end{cases}$	$\begin{cases} 1 \\ k / x \end{cases}$
Cauchy	$\frac{k^2}{2} \log(1 + \frac{x^2}{k^2})$	$x / [1 + x^2/k^2]$	$1 / [1 + x^2/k^2]$

Table 3.3.1: Functions utilized to implement the l^2 -norm, Cauchy, and Huber cost functions, where, where $\rho(x)$ is the robust cost function, $\psi(x) \triangleq \frac{\partial \rho(x)}{\partial x}$ is the influence function, $w(x) \triangleq \frac{\psi(x)}{x}$ is the weighting function, and k is the user defined kernel width.

cise will begin by taking the gradient of the modified cost function (see Eq. 3.10), as presented in Eq. 3.11.

$$\begin{aligned}\frac{\partial J}{\partial X} &= \sum_{n=1}^N \frac{\partial \rho}{\partial e_n} \frac{\partial e_n}{\partial r_n} \frac{\partial r_n}{\partial X} \\ &= r_n(X)^T \left[\frac{\mathbf{1}}{e_n} \frac{\partial \rho}{\partial e} \Big|_{e_n(X)} \Lambda_n^{-1} \right] \frac{\partial r_n}{\partial X}.\end{aligned}\quad (3.11)$$

The expression for the gradient of the modified cost function can be further reduced by noting that $\frac{\mathbf{1}}{e_n} \frac{\partial \rho}{\partial e_n}$, is equivalent to the m-estimator weighting function, by definition. This simplifies the gradient expression for the **m-estimator** cost function as presented in Eq. 3.12, which is equivalent to the traditional l^2 -norm cost function, but with an adaptive covariance. The modified covariance, $\hat{\Lambda}_n$, is provided in Eq. 3.13, where the specific weighting function is dependent upon the utilized m-estimator.

$$\frac{\partial J}{\partial X} = r_n(X)^T \left[w(e_n) \Lambda_n^{-1} \right] \frac{\partial r_n}{\partial X}, \quad (3.12)$$

$$\hat{\Lambda}_n = \left[w(e_n) \Lambda_n^{-1} \right]^{-1} = \frac{\mathbf{1}}{w(e_n)} \Lambda_n, \quad (3.13)$$

From this discussion, it should be noted that the robust state estimation problem when m-estimators are utilized is equivalent to an iteratively re-weighted least squares (IRLS) problem [53, 54]. In IRLS, the effect of an erroneous observation is minimized by covariance adaptation based on the previous optimization iteration's residuals, as presented in Eq. 3.14. The equivalence between m-estimators and IRLS has been previously noted in the literature [55, 56].

$$\hat{X} = \underset{X}{\operatorname{argmin}} \sum_{n=1}^N w_n(e_n) \| r_n(X) \|_{\Lambda_n} \quad (3.14)$$

3.3.2 SWITCH CONSTRAINTS

A more recent advancement within the robotics community, switch constraints (as introduced in [47]) is a specific realization of a lifted state estimation scheme [57]. Switch constraints augment the optimization problem by concurrently solving for the desired states and a set of measurement weighting values, as provided in Eq. 3.15

$$\hat{X}, \hat{S} = \underset{X, S}{\operatorname{argmin}} \sum_{n=1}^N \left[\|\psi(s_n)r_n(X)\|_{\Lambda_n} + \|\gamma_n - s_n\|_{\Xi_n} \right], \quad (3.15)$$

where S is the set of estimated measurement weights, $\psi(\cdot)$ is a real-valued function⁵ such that $\psi(\cdot) \rightarrow [0, 1]$, γ is a prior on the switch constraint, and Ξ is the switch constraint uncertainty model.

To explicitly view the switch constraint technique as a covariance adaptation method, the effect that switchable constraints will have on measurement constraint in the modified cost function (e.g., Eq. 3.15) is evaluated, as depicted in Eq. 3.16 where $w_n \triangleq \psi(s_n)$. This relation can be further reduced to the expression presented in Eq. 3.17, where $\hat{\Lambda}_n^{-1} = w_n^2 \Lambda_n^{-1}$, by noting that w_n is a scalar.

$$\|\psi(s_n)r_n(X)\|_{\Lambda_n} = [w_n r_n(X)]^T \Lambda_n^{-1} [w_n r_n(X)] \quad (3.16)$$

$$= r_n(X)^T \hat{\Lambda}_n^{-1} r_n(X) \quad (3.17)$$

The relation presented in Eq. 3.17 directly shows a relation between the estimated observation weighting — as implemented in the switchable constraint framework — and the scaling of the *a priori* measurement covariance matrix.

⁵Within [15] it is recommended to utilize a piece-wise linear function.

3.3.3 DYNAMIC COVARIANCE SCALING

One potential disadvantage of the switchable constraint framework is the possibility for an increased convergence time due to the augmentation of the optimization space [16]. To combat this issue, the DCS [16] approach was proposed as a closed-form solution to the switch constraint weighting optimization. The m-estimator equivalent weighting function, as derived within [58], is provided as

$$w(x) = \begin{cases} 1, & \text{if } x^2 \leq k. \\ \frac{4k^2}{(x^2+k)^2}, & \text{otherwise,} \end{cases} \quad (3.18)$$

where k is the user defined kernel width, which dictates the residual threshold for observations to be considered erroneous. From this expression, it is noted that DCS is equivalent to the standard least-squares estimator when the residuals are less than the kernel width k ; however, when residuals lie outside of the kernel width, the corresponding observations are de-weighted according to the provided expression.

From the provided weighting expression, the equivalent DCS robust m-estimator can be constructed, as presented in Table 3.3.2. This is performed by utilizing the definitions provided within [55], (i.e., assuming a weighting function is provided, then the influence function is calculated as $\psi(x) = x w(x)$ and the robust cost function is $\rho(x) = \int \psi(x) dx$).

With the DCS equivalent m-estimator, and the discussion provided in section 3.3.1, it can be seen that DCS based state estimation enables robustness through the adaptation of the *a priori* measurement error covariance. The specific relation between the *a priori* and modified covariance matrices is provided in Eq. 3.13 with a weighting function implementation as provided in Eq. 3.18.

M-Estimator	$\rho(x)$	$\psi(x)$	$w(x)$
DCS $\begin{cases} \text{if } x^2 \leq k \\ \text{if } x^2 > k \end{cases}$	$\begin{cases} \frac{1}{2}x^2 \\ \frac{k(3x^2-k)}{2(x^2+k)} \end{cases}$	$\begin{cases} x \\ \frac{4k^2x}{(x^2+k)^2} \end{cases}$	$\begin{cases} 1 \\ \frac{4k^2}{(x^2+k)^2} \end{cases}$

Table 3.3.2: Equivalent dynamic covariance scaling m-estimator, $w(x)$ is the weighting function – as specified in Eq. 3.18, $\psi(X) \triangleq x w(x)$ is the influence function, $\rho(X) = \int \psi(x)dx$ is the robust cost function, and k is the user defined kernel width.

3.3.4 MAX MIXTURES

While the previously discussed estimation frameworks do enable robust inference, they still have the undesirable property that the measurement uncertainty model must adhere to a uni-modal Gaussian model. To relax this assumption, a **GMM** [59] can be utilized to represent the measurement uncertainty model. The **GMM** characterizes the uncertainty model as a weighted sum of Gaussian components, as depicted in Eq. 3.19, where $r_n \triangleq y_n - h_n(X)$ is an observation residual, $w = \{w_1, w_2, \dots, w_m\}$ is the set of mixture weights with the constraint that $\sum_{m=1}^M w_m = 1$, and μ_m, Λ_m are the mixture component mean and covariance, respectively.

$$r_n \sim \sum_{m=1}^M w_m \mathcal{N}(r_n | \mu_m, \Lambda_m) \quad (3.19)$$

Utilizing a **GMM** representation of the measurement uncertainty model can greatly increase the accuracy of the measurement error characterization; however, it also greatly increases the complexity of the optimization problem. This increase in complexity is caused by the inability to reduce the factorization presented in Eq. 3.2 to a **NLLS** formulation, as was possible when a uni-modal Gaussian model was assumed.

To enable the utilization of a **GMM** while minimizing the computational complexity of the opti-

mization process, the max-mixtures approach [17] was proposed. This approach circumvents the increased computational complexity by replacing the summation operation in the **GMM** with the maximum operation, as depicted in Eq. 3.20, where \sim depicts that r_n is approximately distributed according to the right-hand side.

$$r_n \sim \max_m w_m N(r_n | \mu_m, \Lambda_m) \quad (3.20)$$

Within Eq. 3.20, the maximum operator acts as a selector (i.e., for each observation, the maximum operator selects the single Gaussian component from the **GMM** that maximizes the likelihood of the individual observation given the current state estimate). Through this process, each observation is only utilizing the single Gaussian component from the model, which allows for the simplification of the optimization process to the weighted sum of squared residuals [27, 49] (i.e., a **NLLS** optimization problem).

Similar to the previously discussed robust state estimation formulations, the max-mixtures implementation can also be interpreted as enabling robustness through covariance adaptation. Where, the covariance adaptation is enabled through the maximum operator. Specifically, each observation initially belongs to a single component from the **GMM** model, then through the iterative optimization process, the measurement error covariance for each observation is updated to one of the components from within the predefined **GMM**.

3.4 KEY TAKEAWAYS

Utilizing the state estimation approach developed within Section 3.2 (i.e., the **iSAM2** approach) provides an efficient estimation framework when the provided information adheres to the *a pri-*

ori models. However, when the provided information does not adhere to the *a priori* models, the estimator can breakdown [13]. This property is not exclusive to the iSAM2 framework, instead, it is a fundamental property of any estimation framework that exclusively utilizes the l^2 -norm to construct its cost function.

To overcome this limitation, several robust estimation frameworks have been proposed (e.g., m-estimators [14], lifted estimation [47], and MM [17]). Linking all of these estimation frameworks is the concept of enabling robust estimation through appropriately weighting (i.e., scaling the assumed covariance model) the contribution of each information source based upon the level of adherence between the information and the *a priori* model. To implement this concept, the IRLS formation [55], as provided in Eq. 3.14, can be utilized, where the weighting function $w(*)$ is dependent upon the utilized robust estimation framework (i.e., DCS [16]).

At the core of all of these robust frameworks is the assumption that an accurate *a priori* measurement uncertainty model can be provided. As systems become more autonomous, this assumption becomes less valid (i.e., as systems start operating in novel environments, there is no guarantee that the assumed *a priori* measurement uncertainty model characterizes the sensors current observation uncertainty). In an attempt to relax this assumption, a novel robust state estimation framework is developed within Chapter 5, and extended within Chapters 6 and 7.

4

GPS Sensor Model

4.1 CHAPTER ABSTRACT

Within this dissertation, GNSS observations will be the primary information source utilized to validate the proposed robust estimation techniques. Thus, within this chapter, an overview of the utilized GNSS model is provided. To begin the discussion of the utilized GNSS model, it is noted that estimation techniques to precisely localize a kinematic platform with GNSS observables can be broadly partitioned into two categories: differential, or undifferenced. The differential techniques (e.g., real-time kinematic (RTK)) have several attractive properties, such as correlated error mitigation and fast convergence; however, to support a differential processing scheme, an infrastructure of reference stations within a proximity of the platform must be in place to construct observation corrections. This infrastructure requirement makes differential processing techniques infeasible in many locations. To mitigate the need for additional receivers within proximity of the platform, the precise point positioning (PPP) method utilizes accurate orbit and clock models to localize the platform. The autonomy of PPP from local reference stations make it an attractive processing scheme for several applications; however, a current disadvantage of PPP is the slow positioning convergence when compared to differential techniques. In this dissertation, the convergence properties of PPP with an incremental graph optimization scheme (Incremental Smoothing and Mapping (iSAM₂)), which allows for real-time filtering and smoothing, is evaluated. The characterization is first conducted through a Monte Carlo analysis within a simulation environment, which allows for the variations of parameters, such as atmospheric conditions, satellite geometry, and intensity of multipath. Then, an example collected data set is utilized to validate the trends presented in the simulation study.

The software developed for this study can be obtained from the following link: <https://github.com/wv-navLab/PPP-BayesTree>

4.2 LITERATURE REVIEW

As previously mentioned, one commonly used methodology for processing GNSS data is the Precise Point Positioning (PPP) approach [60]. The precise point positioning (PPP) algorithm utilizes the dual-frequency undifferenced GNSS observables, which allows the technique to operate without the need of external reference stations. The undifferenced observations are used along with precise GNSS orbit and clock bias products to mitigate the errors removed through observation differencing [61]. The orbit and clock products that enable the PPP method to achieve decimeter level positioning can be broadcast to an end-user in real-time (e.g., L-band [62] and Iridium modem link [63]).

Real-time kinematic-PPP (kPPP) provides similar positioning performance when compared to traditional differential GPS (DGPS) (i.e., Real Time Kinematic) for dynamic platforms [64]. The comparable positioning accuracy without the need for a nearby static GPS reference station makes it an attractive processing formulation. However, it has been noted in several studies that the PPP formulation has a longer convergence period than comparable differential techniques [65, 66].

In an attempt to decrease the initial convergence period, there has been a plethora of research into augmenting the PPP approach with additional information sources. One of the most commonly utilized augmentation sources for traditional single constellation PPP is additional GNSS observables. One example of this type of augmentation is the incorporation of multiple constellation observations [67, 68]. Another example of this type of augmentation is the PPP-RTK formulation [69] which provides faster convergence by enabling integer ambiguity resolution [70]. Another well studied form of PPP augmentation is the tightly coupled PPP inertial navigation (INS) formulation [71], which has also been shown to decrease the initial convergence period of PPP

[72]. However, all PPP augmented methods require additional infrastructure (e.g., a network of reference stations, or additional sensors on-board), which can be prohibitive for many applications.

Another method to decrease the convergence period of PPP is to utilize a novel optimization framework. This has the potential to provide a benefit over the previously discussed PPP methods because all the previously provided methods utilize the same underlying optimization framework (i.e., a variant of the Kalman filter [9, 11, 30]). The Kalman filter framework estimates the desired states by marginalizing all prior information and propagating with dynamic models to the next time step. For PPP, where a subset of the desired states are not observable over a single epoch (i.e., the carrier-phase ambiguity states), this may not be the best framework.

In this Chapter an evaluation of the PPP model utilizing an incremental graph optimization framework is presented. To enable this evaluation, the advances made within the robotics community on efficient, real-time smoothing are utilized. Research into smoothing, within the robotics community, has been dominated by graph based methodologies since the seminal paper on the subject was published in 1997 [38]. When [38] was published, graph-based smoothing was not widely utilized due to computation complexity of solving the initial formulation. However, quickly thereafter, methods were proposed to greatly reduce complexity through the utilization of factor graphs [73]. The \sqrt{SAM} formulation as presented in [41] was particularly influential as it provided connections between the factor graph formulation and sparse linear algebra. The idea of batch factor graph optimization was later extended to an incremental inference framework in [42, 74]. The work presented in [74] provides a framework to conduct real-time [75], non-linear graph based filter and smoothing.

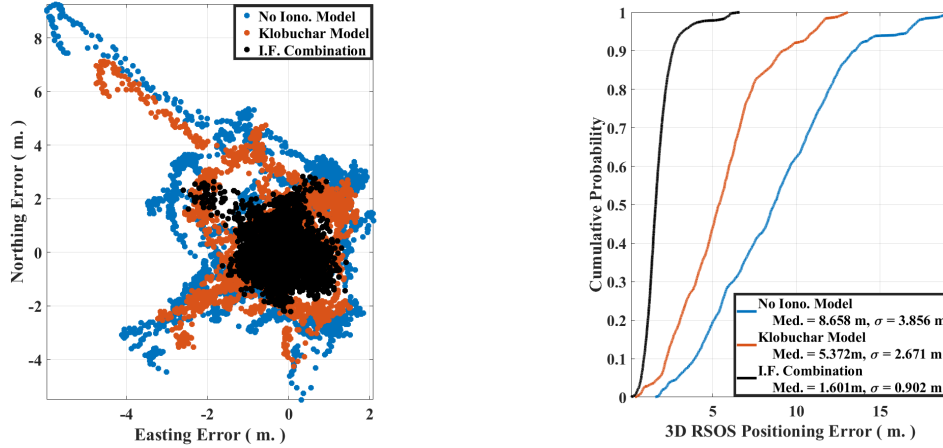
4.3 GPS BASED STATE ESTIMATION

4.3.1 CONSTRUCTING THE FACTOR GRAPH CONSTRAINTS

To allow autonomy of the PPP approach from local reference stations, the undifferenced dual-frequency GNSS observables are utilized. Due to the undifferenced nature of the observations, the PPP processing technique must incorporate GNSS error mitigation models — these models provide corrections for the corrupting sources that would be mitigated through observation differencing — to provide an accurate positioning solution. The sources that corrupt a GNSS observation can be segregated into three partitions: the error contributed by the propagation medium, the error contributed by the control segment, and the error contributed by the user.

To begin constructing our measurement model, the method implemented to mitigate the propagation medium errors are discussed. The error attributed to the propagation medium is composed of delay due to the ionosphere and the delay due to the troposphere. To mitigate the ionospheric delay, it is common to leverage the dispersive nature of the medium, and a linear combination of the GPS L_1 and L_2 frequencies (1575.42 MHz and 1227.60 MHz, respectively) is formed to produce ionospheric-free (IF) pseudorange and carrier phase measurements [61]. The IF combination of an observable, O^j , can be seen in Eq. 4.1. However, if only single frequency observations are available, an ionospheric model (i.e., the Klobuchar model [76]) can be utilized to reduce the error attributed to the medium. To see the affect that the selected ionosphere model can have on the localization performance, Fig 4.3.1 is provided.

$$O_{IF}^j = O_{L_1}^j \left[\frac{f_1}{f_1^2 - f_2^2} \right] - O_{L_2}^j \left[\frac{f_2}{f_1^2 - f_2^2} \right] \quad (4.1)$$

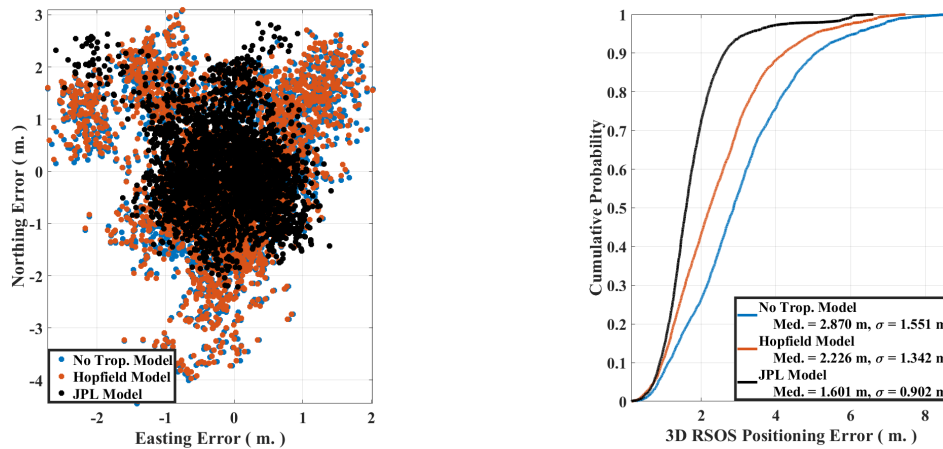


(a) Horizontal localization error as a function of the ionosphere model. (b) Total localization (i.e., 3D) error as a function of the ionosphere model.

Figure 4.3.1: GNSS based localization error as a function of the utilized ionosphere model.

To mitigate the error attributed to the troposphere, both the wet and the dry component of the troposphere must be modeled, as shown in Eq. 4.2. One commonly utilized model for the dry component of the troposphere is the Hopfield approach [77]. To compensate for the wet delay — the wet component only accounts for approximately 10% of the total troposphere error — and the residual error of the dry delay model, a stochastic random variable is added to the state vector. Additionally, there are also higher fidelity empirical models (i.e. the JPL model [77]) that can be broadcast to the user. To evaluate the affect that the selected troposphere model can have on the localization performance, Fig 4.3.2 is provided.

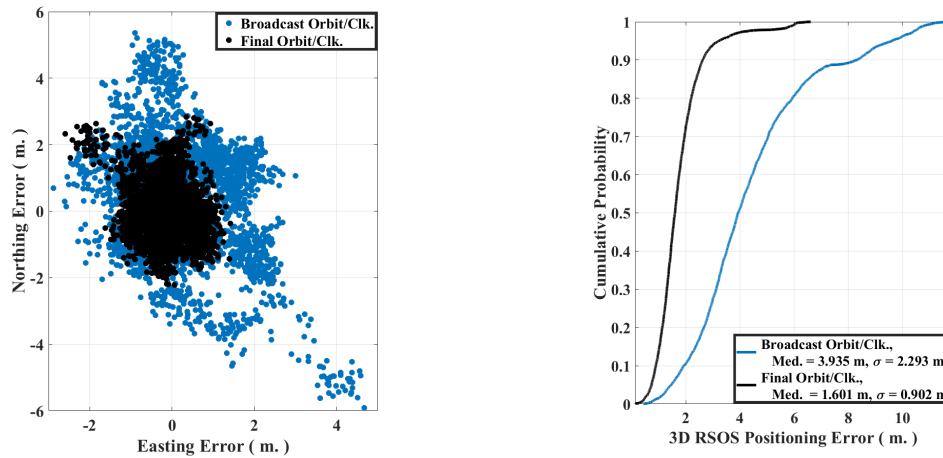
$$T(el) = T_{z,d}M_d(el) + T_{z,w}M_w(el) \quad (4.2)$$



(a) Horizontal localization error as a function of the troposphere model. (b) Total localization (i.e., 3D) error as a function of the troposphere model.

Figure 4.3.2: GNSS based localization error as a function of the utilized troposphere model.

To mitigate the error attributed to the control segment, the PPP approach utilizes orbit and clock corrections. These global corrections are generated through a network of reference stations that are at a set of known locations, and thus can be utilized to accurately estimate the satellites states. The orbit and clock products that enable the PPP method to achieve decimeter level positioning can be broadcast to an end-user in real-time (e.g., L-band [62] and Iridium modem link [63]). To evaluate the affect that the selected orbit and clock model can have on the localization performance, Fig 4.3.3 is provided. For a more thorough review of the affect that the quality of the orbit and clock models can have on the localization performance of GNSS based estimation, the reader is referred to [34, 72].



(a) Horizontal localization error as a function of the orbit and clock model. (b) Total localization (i.e., 3D) error as a function of the orbit and clock model

Figure 4.3.3: GNSS based localization error as a function of the utilized satellite orbit and clock model.

Finally, a discussion on the user error segment is provided. The user error segment is composed of two sources: multipath error, and receiver thermal noise error. For this study, no methods were implemented to explicitly model the user segment error; however, as noted in [78], the magnitude of the user error is proportional to the elevation angle between the platform and the satellite so, within this evaluation, the uncertainty in the observation is scaled by the elevation angle. It should be noted that PPP observational models for moving platforms typically include corrections for relativistic effects (i.e., from the GPS broadcast correction), receiver and satellite antenna phase center variation, and carrier-phase wind-up; however, these effects were neglected within this simulation study. Additionally, dynamic platform generally couple inertial information with the GNSS observables to mitigate uncertainty in the platforms dynamic model [72].

Utilizing the provided error mitigation techniques, the PPP observation model can be constructed. The pseudorange and carrier-phase measurements are modeled as shown in Eq. 4.3 and Eq. 4.4, re-

spectively: where, δt_u is the receiver's clock bias, δt_s is the satellite's clock bias, T_z is the tropospheric delay in the zenith direction, $M_d(e^{\ell'})$ is a user to satellite elevation angle dependent mapping function, $\delta_{Rel.}$ is the correction attributed to relativistic effect [79], $\delta_{P.C.}$ is the correction attributed to the offset between the satellite's center of mass and the phase center of the antenna [80], $\delta_{D.C.B}$ is the differential code bias correction [77], $\delta_{W.U.}$ is the correction attributed to the windup effect on the phase observables [81], λ_{IF} is the wavelength corresponding to the IF combination, and N_{IF} is phase ambiguity. In Eqs. 4.3 and 4.4 the remaining unmodelled error sources are indicated with ε . To implement the provided observation model in software, the open-source library GPSTk [82] is utilized.

$$\begin{aligned} \rho_{IF}^j = & \|X_s^j(t_e) - X_u^j(t_r)\| + c(\delta t_u - \delta t_s) \\ & + T_{z,d}M_d(e^{\ell'}) + \delta_{Rel.} + \delta_{P.C.} + \delta_{D.C.B} + \varepsilon_{\rho}^j \end{aligned} \quad (4.3)$$

$$\begin{aligned} \varphi_{IF}^j = & \|X_s^j(t_e) - X_u^j(t_r)\| + c(\delta t_u - \delta t_s) \\ & + T_{z,d}M_d(e^{\ell'}) + \delta_{Rel.} + \delta_{P.C.} + \delta_{W.U.} + \lambda_{IF}N_{IF}^j + \varepsilon_{\varphi}^j \end{aligned} \quad (4.4)$$

Using the PPP observation model, a GNSS constraint for the factor graph [83] can be constructed. To begin, it is noted that the GNSS observations are providing a set of likelihood constraints, $L(O|X)$, on the optimization process. If the assumption is made that the state and measurement noise models are Gaussian, then this constraint can be incorporated into factor graph through the mahalanobis distance, as provided in Eq. 4.5, where z is the observed measurement, \hat{z} is the estimated measurement — calculated using Eq's 4.3 and 4.4 — and Λ represents the uncer-

tainty of the observation.

$$\begin{aligned}\psi_l &= \|z - \hat{z}\|_\Lambda \\ &= (z - \hat{z})^T \Lambda^{-1} (z - \hat{z})\end{aligned}\tag{4.5}$$

INCORPORATING THE CARRIER-PHASE AMBIGUITY STATE

There are several ways in which the carrier-phase ambiguity states can be incorporated into the factor graph. One such way is to incorporate a new carrier-phase ambiguity state for each epoch. Consecutive carrier-phase ambiguity states can be constrained by a process noise update. The measurement Jacobian associated with this graph construction is represented in Fig. 4.3.4.A. From Fig 4.3.4.A, it can be seen that the measurement Jacobian is more densely populated than desired.

To construct less densely populated measurement Jacobian (i.e., a more efficient optimization scheme), the knowledge that the true carrier-phase ambiguity value for a given satellite within a continually tracked phase-arc is a constant value can be leveraged. Due to this property of the true ambiguity value, the carrier-phase ambiguity factor can be represented as a random constant variable. Where, initially, a single factor is added for each satellite, and a new factor is added only if there is a cycle-slip or if a new satellite is tracked. By treating the carrier-phase bias factor in this manner (i.e., like a “landmark” variable in traditional pose-graph SLAM [36]), and utilizing the Bayes tree based optimizer, an efficient real-time smoothing formulation for GNSS signal processing is presented. The measurement Jacobian associated with this graph construction is represented in Fig. 4.3.4.B. From Fig. 4.3.4.B, a less densely populated measurement Jacobian is seen, as desired.

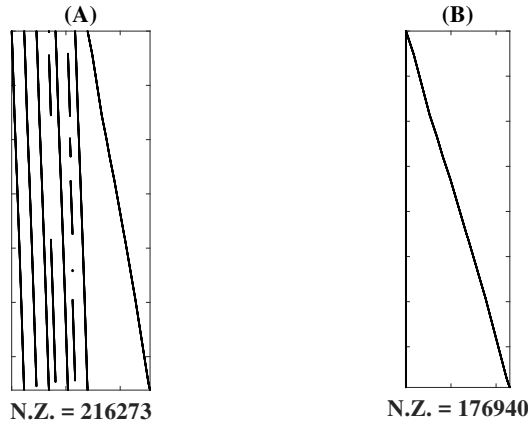


Figure 4.3.4: Example sparse measurement Jacobian for the PPP processing strategy. Figure (A) shows the measurement Jacobian when a new carrier-phase ambiguity state is added for each epoch, where N.Z. represents the number of non-zero elements. Figure (B) shows the measurement Jacobian when a new carrier-phase ambiguity is added only when a new satellite is tracked or if a carrier-phase cycle-slip occurs, where N.Z. is the number of non-zero elements.

4.3.2 INCREMENTAL FACTOR GRAPH INFERENCE

The formulation presented in the previous sections provides an efficient estimation of \hat{X} when all of the information is provided *a priori*. However, it is generally the case that information is arriving sequentially, and it is desired to incrementally provide state estimates. The ability to provide an incremental estimator lies in the capability of the optimizer to reuse prior computations. A well studied technique of computation reuse, for state estimation, is to employ QR-factorization to update the previous matrix factorization [30, 42]; however, this technique only works for linearized systems.

To overcome this limitation, the Incremental Smoothing and Mapping (iSAM2) formulation was developed [74]. The iSAM2 formulation allows for incremental inference over linear or non-linear objective functions through the utilization of a novel graphical model, the Bayes tree [84].

To provide insight into this formulation, specifically for GNSS applications, a simple GNSS example will be presented. Where it will be shown how to convert the GNSS factor graph into a Bayes tree. Additionally, a discussion will be provided on how the Bayes tree graphical models allows for efficient inference.

To begin our discussion, a factor graph that represents the GNSS inference problem is presented in Fig. 4.3.5. With this factor graph, it is desired to estimate the states $\{X, B\}$. In this formulation, X represents the position, troposphere, and receiver clock bias states, as provided in Eq. 4.6. Additionally, the vertices B represents the carrier-phase bias states.

$$X = \begin{pmatrix} \delta P \\ T_{z,w} \\ C_b \end{pmatrix} \quad (4.6)$$

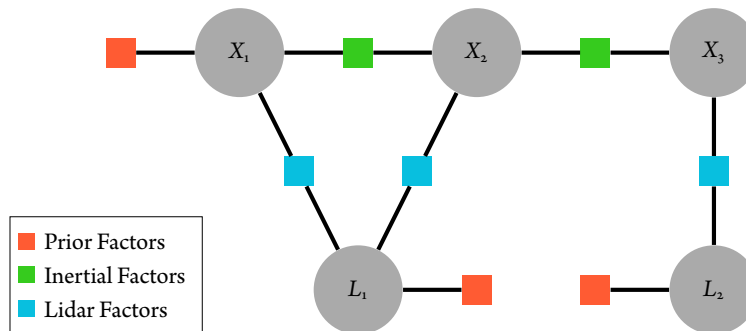


Figure 4.3.5: GNSS factor graph construction where the carrier-phase ambiguity states are treated as “land-mark” observations.

Utilizing the factor graph presented in Fig. 4.3.5, the process of converting a factor graph into a Bayes tree can commence. To enable this, an intermediate step an interment step is taken to construct a Bayes net. The Bayes net can be constructed from the factor graph using a variable

elimination game [85]. For our specific example, the Bayes net is provided in Fig. 4.3.6, if the bias states are eliminated first then the positing states (i.e., the variable elimination is B_1, B_2, X_1, X_2, X_3). It should be noted that if the elimination ordering is varied, the resultant Bayes net will change and this can have a substantial impact on the run-time of the optimizer [86].

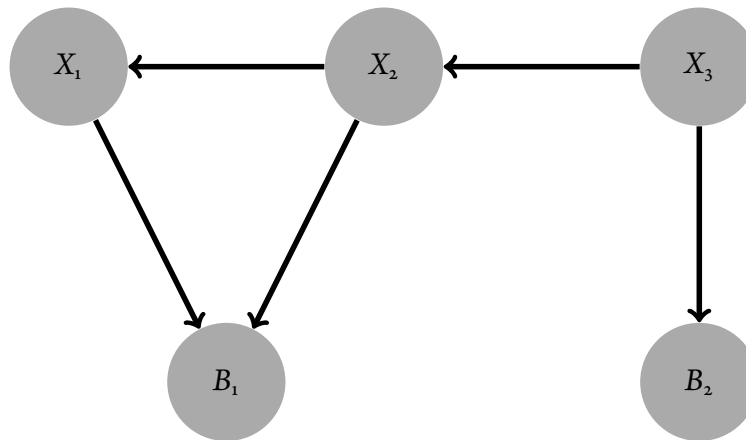


Figure 4.3.6: Generating a Bayes Net from the original factor graph, as presented within Fig. 4.3.5, using the elimination ordering, $\{B_1, B_2, X_1, X_2, X_3\}$

Utilizing the previously constructed Bayes net, the Bayes tree can be constructed. The Bayes tree is constructed to take advantage of the clique structure within the Bayes net. That is, by re-writing the Bayes net the resulting structure is a directed tree, where the vertices in the graph represent cliques in the original Bayes net. For our GNSS example, the constructed Bayes tree is provided in Fig. 4.3.7.

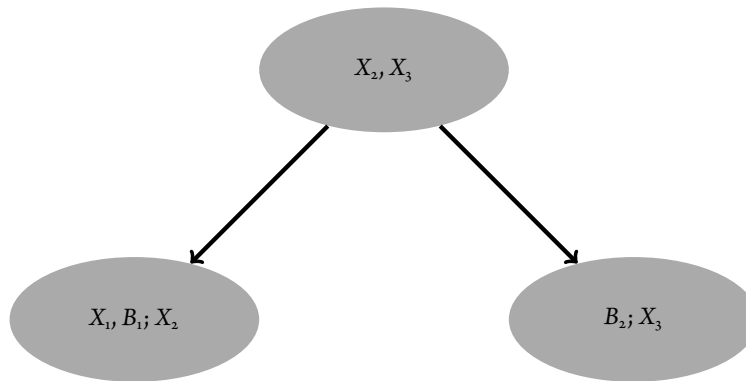


Figure 4.3.7: Generating the Bayes tree from cliques in the chordal Bayes net, as constructed within Fig. 4.3.6.

The tree structure present in the Bayes tree plays a pivotal role in the ability of the data structure to provide an efficient incremental inference engine. The primary advantage of the tree structure is in the idea that only local sections (i.e., a branch in the Bayes tree structure) of the data structure needs to be re-linearized when new constraints are added to the graph. Where re-linearization can be conducted by converting a subset of the Bayes tree back into a factor graph adding the new constraint [27].

4.4 EXPERIMENTAL EVALUATION

To conduct an evaluation of the proposed GNSS processing strategy, two data sets will be utilized. First, the evaluation is conducted within a simulation environment. Then, to provide a more diverse evaluation, the proposed approach is tested on GNSS data collected on-board a UAV.

4.4.1 EVALUATION ON SIMULATED DATA

To conduct an analysis of PPP convergence, a simulation environment was constructed. For synthetic observations generation, the SatNav-3.04 Toolbox [87] is utilized, which provides a Matlab environment for generating dual-frequency pseudorange and carrier-phase observations for a specified trajectory. For this evaluation, four trajectories of varying dynamic were created — an example flight trajectory is provided in Fig. 4.4.1. To tailor the toolbox for an evaluation of kinematic PPP airborne positioning, several minor modifications were made, as discussed in [34]. For example, toolbox was modified to include attitude dependent satellite masking and carrier-phase breaks. That is, when a satellite is obscured or nearly obscured due to a change in platform attitude, it is masked from view and the the potential of a carrier-phase breaks is increased. Additionally, for a PPP analysis, a methodology of constructing an orbit and clock model is required. This error model is constructed by differencing JPL's International GNSS Service (IGS) submission with European Center for Orbit Determination (CODE) submission. For a more detailed discussion on the simulation environment, the reader is directed to [34].

To evaluate the positioning performance of the PPP incremental graph optimizer, a Monte Carlo style experiment was implemented. Specifically, one hundred datasets were generated where several parameters, which are known to adversely affect GNSS positioning performance, were ran-

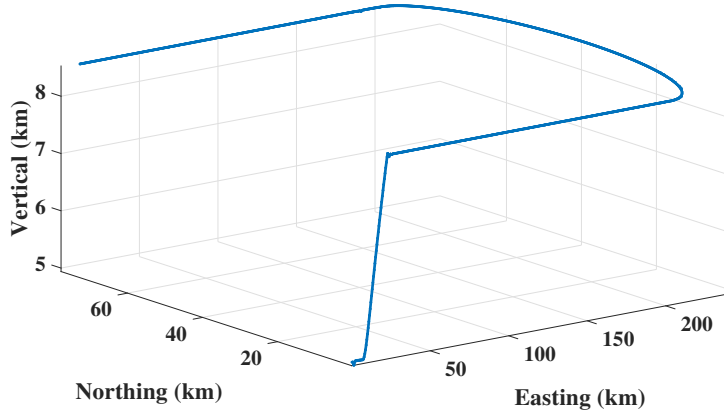


Figure 4.4.1: Example flight trajectory utilized for simulated GNSS observation generation.

domly initialized for each flight — see Table 4.4.1 for additional information on varied parameters.

Table 4.4.1: Description of randomly initialized parameters within the Monte-Carlo evaluation.

Varied parameters	Parameter description
Thermal noise	$\sigma_\rho = 0.32m, \sigma_\phi = 0.16\lambda$
Multipath	$\sigma = 0.4m, \tau = 15sec$
Tropospheric delay	Modified Hopfield with linear scale
Ionospheric delay	O_{IF} used to mitigate error to 1 st order
Receiver clock bias	$\sigma = 30ns, \delta\tau_b = 100ns$
Phase ambiguity	Random initialization with attitude dependent phase breaks
Orbits	Orbits $\sigma = 5cm$ with linear scale

To provide a reference positioning solution, a traditional extended Kalman filter (EKF) was utilized, where the specific EKF formulation details are provided in [34]. To provide a fair comparison, the same stochastic models were implemented for both estimators. The specific stochastic models utilized for the comparison are provided in Table 4.4.2.

Table 4.4.2: Stochastic model parameters for estimators.

Parameter	<i>a priori</i> σ	Process noise	Correlation time
Position	1.0 m	5 $\frac{m}{\sqrt{s}}$	∞
Trop. wet zenith delay	0.3 m	3e-5 $\frac{m}{\sqrt{s}}$	∞
Receiver clock	3e6 m	2000 $\frac{m}{\sqrt{s}}$	0
Phase biases	100 m	0 $\frac{m}{\sqrt{s}}$	∞

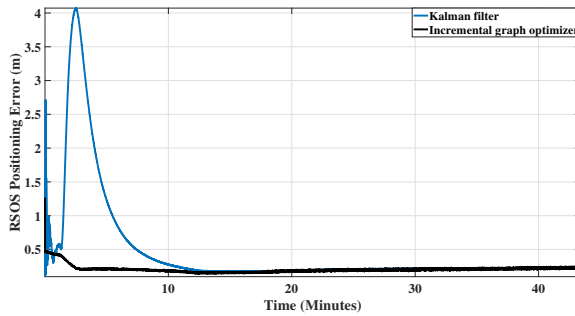


Figure 4.4.2: Example RSOS positioning error profile for a typical simulated data set

Table 4.4.3: Positioning statistics for a single flight

(cm.)	Incremental graph	Kalman filter
Median	20.09	21.40
Mean	20.56	45.68
Variance	5.13	72.23
Max.	126.7	407.26

ESTIMATION PERFORMANCE

To begin an evaluation of the PPP incremental graph optimizer, a single data set, which is representative of all datasets simulated for this study, will be analyzed. As a starting point, the residual sum of squares (RSOS) positioning error is utilized to evaluate the performance of both estimators, as shown in Fig. 4.4.2. From Fig. 4.4.2, it can be readily seen that the incremental graph optimizer more quickly converges — when compared to the EKF — to a steady-state value. Additionally, it should be noted that both estimators converge to approximately the same value.

The RSOS positioning error statistics for both estimators are provided in Table 4.4.3. From Table 4.4.3 it should be noted that the incremental graph optimizer outperforms the EKF with respect to all metrics provided (e.g., the incremental graph optimizer provides a 25 cm error reduction with

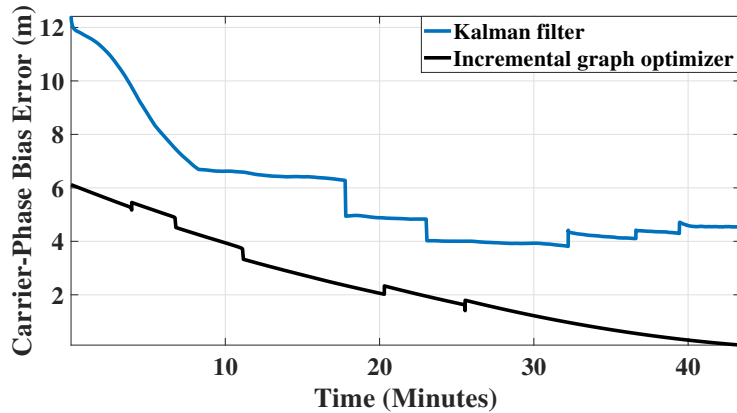


Figure 4.4.3: Example phase bias convergence rate for a typical simulated data set

respect to the mean RSOS positioning error).

To continue an analysis of this example data set, it can be seen in Fig. 4.4.2 that the most substantial positioning error reduction attributed to the incremental graph optimizer occurs during the first several minutes of the flight (i.e., during the PPP convergence period). To provide insight into the accelerated convergence rate of the incremental optimizer, next, an evaluation of both estimators ability to correctly resolve the phase bias states is provided in Fig. 4.4.3. Where it can be seen that the incremental graph optimizer provides a substantial decrease in carrier-phase bias estimation error when compared to the EKF.

Now, the evaluation shifts from a single flight to the performance of both estimators over all simulated data sets. As with the previous evaluation of a single flight, the RSOS positioning error will be utilized as the metric of comparison.

To begin our evaluation, the cumulative distribution function (CDF) of the RSOS positioning error for both estimators is evaluated, as provided in Fig. 4.4.4. From Fig. 4.4.4 it can be noted that there is a considerable shift to the left for the CDF of the RSOS positioning error of the incremental

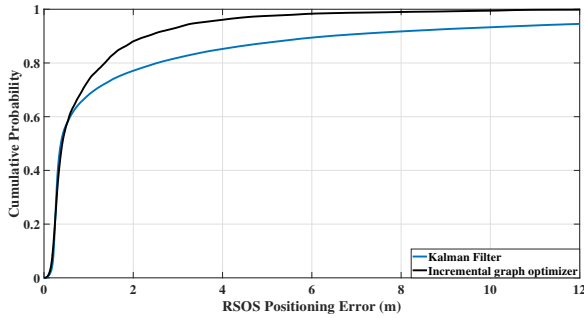


Table 4.4.4: Positioning statistics for all epochs.

(cm.)	Incremental graph	Kalman filter
Median	40.01	39.7
Mean	97.50	256.27
Variance	149.46	641.40
Max.	2318.43	10,152.28

Figure 4.4.4: CDF of the RSOS positioning error for all epochs over the 100 simulated datasets.

graph optimizer for large error values. One possible explanation for this trend — as indicated by our evaluation of a single flight — is that the incremental graph optimizer is more quickly converging to a steady-state value.

To confirm that RSOS positioning error seen in Fig. 4.4.4 for the incremental graph optimizer is occurring during the initial convergence period, a CDF of the RSOS positioning error for both optimizers during the first 15 minutes of each data set is provided in Fig. 4.4.5. As indicated by the right shift in Fig. 4.4.5 of the EKF RSOS positioning error line, the incremental graph optimizer provides a more accurate positioning solution during the initial convergence period. The specific RSOS positioning error reduction during the initial convergence period can be seen in Table 4.4.5.

4.4.2 EVALUATION ON COLLECTED DATA

EXPERIMENTAL SETUP

Finally, to verify the positioning performance benefits noted in the simulation study, a similar analysis is conducted on an example collected dataset. The dataset to be evaluated was collected on-

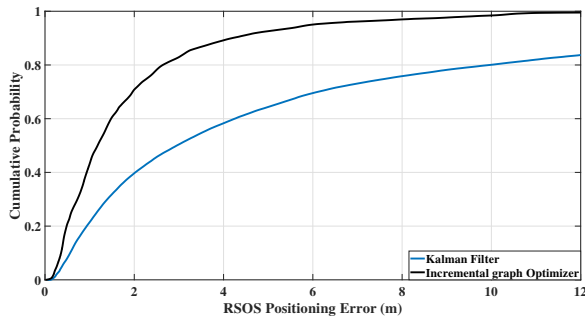


Table 4.4.5: Positioning statistics for epochs within the convergence period.

(cm.)	Incremental Graph	Kalman filter
Median	117.76	295.15
Mean	188.79	665.35
Variance	217.16	984.01
Max.	2,318.43	10,152.28

Figure 4.4.5: CDF of the RSOS positioning error for epochs within the convergence period over the 100 simulated datasets.



Figure 4.4.6: *Phastball* research platform [1] in flight over the West Virginia University Jackson's Mill airfield.

board a small, fixed-wing Unmanned Aerial Vehicle (UAV). This UAV testbed (*Phastball*) — as depicted in Fig. 4.4.6 — was developed at West Virginia University as a research platform [1].

The *Phastball* is equipped with a NovAtel OEM-615 dual-frequency GNSS receiver, which provides 10 Hz GNSS observables over the duration of the flight. The flight profile is depicted in Fig. 4.4.7. A second OEM-615 NovAtel GNSS receiver was placed near the runway to allow for a post-processed RTK solution, where RTKLIB [88] was utilized to generated the reference solution.

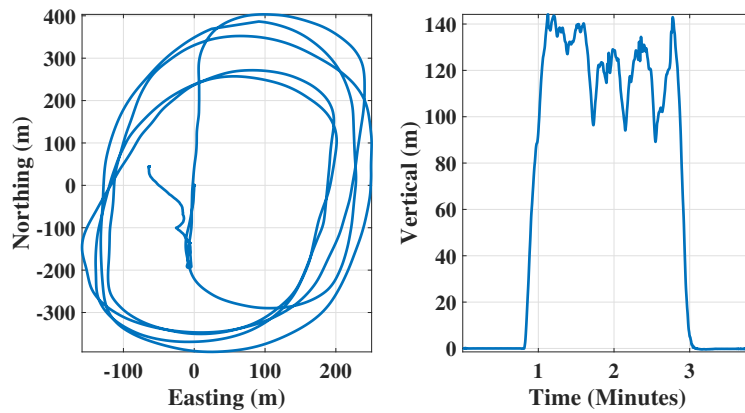


Figure 4.4.7: Flight profile for collected data set.

ESTIMATION PERFORMANCE

Utilizing this dataset, the PPP incremental graph optimizer is evaluated against a Kalman filter based PPP approach, where both estimators are given the same initial conditions (i.e., both estimators are provided the same initial error covariance, the same measurement noise model, and the same process noise model). In Fig. 4.4.8 the 3D RSOS positioning error for both estimators is provided. The result presented in Fig. 4.4.8 follow the trend provided by the simulation study (i.e., the incremental graph optimizer provides faster positioning error convergence than that provided by the Kalman filter). This is further validated by looking at Table 4.4.6, where a substantial median 3D RSOS positioning error reduction is granted by the PPP incremental graph optimizer when compared to the Kalman filter.

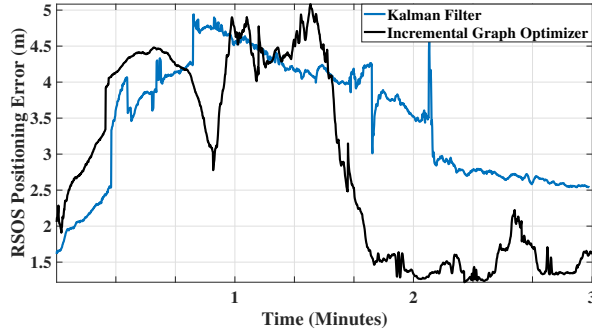


Table 4.4.6: RSOS localization statistics for collected dataset.

(cm.)	Incremental graph	Kalman filter
Median	180.23	373.56
Mean	260.76	349.08
Variance	127.82	84.06
Max.	508.32	494.07

Figure 4.4.8: RSOS positioning error comparison for a Kalman filter and the incremental graph optimizer when a Phastball collected dataset is utilized.

4.5 KEY TAKEAWAY

The desire to precisely localize a platform is of paramount importance to a myriad of fields. This desire has led to a plethora of research into precise GNSS localization due to its ability to provide a precise and globally consistent solution. One of the most commonly utilized GNSS formulations is the precise point positioning (PPP) technique due to its autonomy from local reference stations. However, it has been noted in several studies that PPP has a relatively long initial convergence period when compared to differential techniques.

To reduce the convergence time of PPP, this chapter proposes the use of recent advances in real-time smoothing made within the robotics community. Specifically, this chapter makes connections between GNSS localization and incremental pose-graph optimization. The connection between the two fields lies in the ability to treat phase bias states as “landmark” nodes in the graph. By treating the phase bias state in this manner, and utilizing a Bayes tree based optimizer, efficient smoothing of the position states can be conducted in real-time.

To quantify the benefit of this formulation, a Monte-Carlo style experiment was conducted

within a simulation environment. Utilizing the simulated data, the incremental graph optimization was evaluated along with a traditional [EKF-PPP](#) formulation. Through this evaluation, it was found that the incremental graph optimization technique provided a substantial RSOS positioning error reduction during the initial [PPP](#) convergence period when compared to a traditional [EKF](#) formulation. This finding was also validated with an evaluation of a short duration dataset collected with a fixed-wing UAV.

5

Batch Covariance Estimation (BCE)

5.1 CHAPTER ABSTRACT

As robotic systems become more ubiquitous, it is necessary to develop robust state estimation algorithms that are able to withstand novel and non-cooperative environments. When dealing with novel and non-cooperative environments, little is known *a priori* about the measurement error uncertainty, thus, there is a requirement that the uncertainty models of the localization algorithm be adaptive. Within this chapter, the batch covariance estimation technique, which enables robust state estimation through the iterative adaptation of the measurement uncertainty model, is proposed. The adaptation of the measurement uncertainty model is granted through non-parametric clustering of the residuals, which enables the characterization of the measurement uncertainty via a Gaussian mixture model. The provided Gaussian mixture model can be utilized within any non-linear least squares optimization algorithm by approximately characterizing each observation with the sufficient statistics of the assigned cluster (i.e., each observation's uncertainty model is updated based upon the assignment provided by the non-parametric clustering algorithm). The proposed algorithm is verified on several GNSS collected data sets, where it is shown that the proposed technique exhibits some advantages when compared to other robust estimation techniques when confronted with degraded data quality.

5.2 BCE ALGORITHM FORMULATION

As shown by the discussion in Chapter 3, several robust state estimation frameworks have been proposed. Each of these methodologies work under the assumption that an accurate *a priori* measure-

All developed software and data utilized within this chapter is publicly available at <https://bit.ly/2V4QuBo>.

ment error covariance is available. However, in practice, the requirement to supply an accurate *a priori* characterization of the measurement uncertainty model is not always feasible when considering that the platform could be operating in a novel environment¹, a non-cooperative environment², or both.

Within this chapter, a novel framework is proposed that addresses the issue of robust state estimation where, the framework is not only robust to erroneous observations, but also erroneous *a priori* measurement uncertainty estimates. To elaborate on the proposed framework, the remainder of this chapter proceeds in the following manner: first, the assumed data model is discussed; then, a method for learning the measurement error uncertainty model from residuals is provided; finally, the proposed robust estimation framework is described.

5.2.1 DATA MODEL

Given a set of residuals $R = \{r_1, r_2, \dots, r_n\}$ with $r_n = y_n - h_n(X) \in \mathbb{R}^d$. It is assumed that the set can be partitioned into groupings of similar residuals (i.e., $\bigcup_{m=1}^M C_m = R$), where each group, C_k , can be fully characterized by a Gaussian distribution (i.e., $C_k \sim N(\mu_k, \Lambda_k)$). Given this assumptions, the set of residuals are fully characterized as a weighted sum of Gaussian distributions (i.e., a GMM), as depicted in Eq. 5.1

$$r_n \sim \sum_{m=1}^M w_m N(r_n | \theta_m) \quad \text{s.t.} \quad \theta_m \triangleq \{\mu_m, \Lambda_m\}, \quad (5.1)$$

where, w is the set of mixture weights with the constraint that $\sum_m w_m = 1$, and θ_m is the mixture

¹A novel environment is simply one that has not before been previously experienced.

²A non-cooperative environment is one that emits characteristics that inhibit the sensors ability to provide accurate observations.

components sufficient statistics (i.e., the mixture component mean and covariance).

To enable the explicit assignment of each data point to a Gaussian component within the GMM, an additional latent parameter, $Z = \{z_1, z_2, \dots, z_n\}$ with $z_n \in \mathbb{R}^M$, is incorporated. The assignment variables Z are characterized according to a categorical distribution, as depicted in Eq. 5.2, where w_m is the component weight, and $I_{[*]}$ is the indicator function which evaluates to 1 if the equality within the brackets is true and 0 otherwise.

$$p(z_n | w) = \prod_{m=1}^M w_m I_{[z_n=m]} \quad (5.2)$$

Through the incorporation of the categorically distributed assignment parameters into the GMM, every data instance $r_n \in R$ can be characterized through Eq. 5.3. Where, this equation explicitly encodes that each data instance is characterized by a single Gaussian component from the full GMM.

$$p(r_n | \theta_m, z_m) = \sum_{m=1}^M w_m \mathcal{N}(r_n | \theta_m) I_{[z_n=m]} \quad (5.3)$$

As a Bayesian framework will be utilized for model fitting (as described in section 5.2.2), a prior distribution must be defined for the GMM mixture weights w , and the sufficient statistics, θ . To define the prior over the mixture weights, w , a Dirichlet distribution [89] is utilized, as it is a conjugate prior³ to the categorical distribution [91]. The specific form of the Dirichlet prior is provided in Eq. 5.4, where $\alpha = (\alpha_1, \dots, \alpha_k)$ ⁴ is a set of hyper-parameters such that $\alpha_i > 0$, and S is the

³Conjugate priors are commonly utilized because they make the calculation of posterior distribution tractable through the *a priori* knowledge of the posterior distribution family [90].

⁴For the analysis presented within this study, a symmetric Dirichlet prior was utilized (i.e., $\alpha_k = 1 \forall k$).

probability simplex⁵.

$$p(w|\alpha) = \frac{\Gamma(\sum_{m=1}^M \alpha_m)}{\prod_{m=1}^M \Gamma(\alpha_m)} \prod_{m=1}^M w_m^{\alpha_m-1} \mathbf{1}_{[w_m \in S]} \quad (5.4)$$

To define a prior over the GMM sufficient statistics, θ , the normal inverse Wishart (NIW) is utilized due to the conjugate relation with the multivariate Gaussian that has unknown mean and covariance [92]. The NIW distribution is defined in Eq. 5.5, where $\beta = \{m_o, \kappa_o, \nu_o, S_o\}$ ⁶ is the set of hyper-parameters, and W^{-1} is the inverse Wishart distribution, as discussed within [93], where $\Gamma_d(\ast)$ is the multivariate gamma function, $\text{Tr}(\ast)$ is the matrix trace, and d is the dimension of S_o .

$$p(\theta | \beta) = N(\mu | m_o, \frac{1}{\kappa_o} \Lambda) W^{-1}(\Lambda | S_o, \nu_o) \quad (5.5)$$

$$\text{s.t. } W^{-1}(\Lambda | S_o, \nu_o) = \frac{|\Lambda|^{\frac{\nu_o-d-1}{2}}}{2^{\nu_o \frac{d}{2}} \Gamma_d(\frac{\nu_o}{2}) |S_o|^{\frac{\nu_o}{2}}} e^{-\frac{1}{2} \text{Tr}(S_o^{-1} \Lambda)}$$

Utilizing the data model defined earlier, the joint probability distribution that characterizes the system is defined in Eq. 5.6. Additionally, a visual representation of the underlying model is graphically represented in Fig. 5.2.1.

⁵A simplex is simply a generalization of the triangle to n-dimensional space (i.e., $S = \{w \in \mathbb{R}^m : w_i \geq 0 : \sum_m w_m = 1\}$).

⁶The NIW hyper-parameters can be intuitively understood as follows: m_o is the expected prior on the mean vector μ with uncertainty κ_o , and ν_o is the expected prior on the covariance matrix Λ with uncertainty S_o

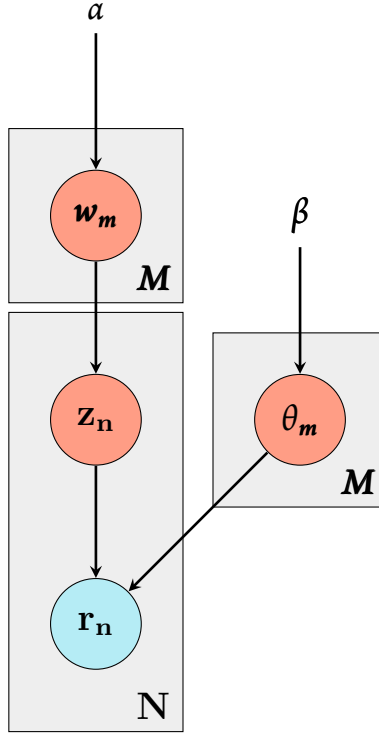


Figure 5.2.1: Graphical representation of the utilized data model which visually encodes the joint probability distribution presented in Eq. 5.6. Each red node represents a latent parameter, and the blue node represents the provided observations. The plate notation [2] is utilized to represent the replication of a random variable.

$$\begin{aligned}
 p(R, \theta, Z) &= p(w | a) \cdot \\
 &\prod_{n=1}^N p(z_n | w_n) p(r_n | z_n, \theta) \cdot \\
 &\prod_{m=1}^M p(\theta_m | \beta)
 \end{aligned} \tag{5.6}$$

5.2.2 GAUSSIAN MIXTURE MODEL FITTING

Provided with the joint distribution, as depicted in Eq. 5.6, the ultimate goal of a GMM fitting algorithm is to estimate the model parameters that maximize the log marginal likelihood, as provided in Eq. 7.3, of the observations [59, 94]. Due to the dimensionality of the space, the integral presented in Eq. 7.3 is not computational tractable in general [95]. Thus, techniques for approximating the integral need to be utilized.

$$\log p(R) = \log \int p(R, \theta, Z) d\mathbf{Z} d\theta \quad (5.7)$$

In practice, two broad classes of integral approximation algorithms are utilized: Monte Carlo methods [96], and variational methods [97]. The Monte Carlo methods approximate the integral by averaging the response of the integrand for a finite number of samples⁷. In contrast, the variational approaches convert from an integration into an optimization problem, by defining a family of simplified functions (i.e., the family of exponential probability distributions) and optimizing the function within that family that most closely matches⁸ the original function [97].

Within this work, it was elected to utilize a variational model fitting approach, as this class of algorithms dramatically reduce the computational complexity of the problem when compared to Monte Carlo techniques [95]. To implement this approach, for the proposed data model, the assumption is made that the joint distribution defined in Eq. 5.6 can be represented as, $p(R, \theta, Z) \approx q(\theta) q(Z)$, which is commonly referred to as the mean-field approximation [97]. Utilizing the

⁷Considerable research is dedicated to efficient and intelligent ways to sample a space. For a succinct review of Monte Carlo sampling approaches, the reader is directed to [96].

⁸Closeness in this context is usually measured with the Kullback-Leibler divergence [95, 97] or free-energy [94].

mean-field approximation, the log marginal likelihood is represented as Eq. 5.8. Through the application of Jensens's inequality [98], the equality in Eq. 5.8 can be converted into a lower bound on the log marginal likelihood⁹, as presented in Eq. 5.9.

$$\log p(R) = \log \int q(Z) q(\theta) \frac{p(R, \theta, Z)}{q(Z) q(\theta)} dZ d\theta \quad (5.8)$$

$$\begin{aligned} &\geq \int q(Z) q(\theta) \log \frac{p(R, Z | \theta)}{q(Z)} dZ d\theta + \\ &\int q(\theta) \log \frac{p(\theta)}{q(\theta)} d\theta \end{aligned} \quad (5.9)$$

The right hand side of the inequality presented in Eq. 5.9 is commonly referred as the free energy functional [94], as depicted in Eq. 5.10. As it is a lower bound on the log marginal likelihood, an optimal set of model parameters can be found through iterative optimization.

$$\log p(R) \geq F [q(Z), q(\theta)] \quad (5.10)$$

To find the set of equations necessary for iterative parameter optimization, the partial derivative of the free energy function can be taken with respect to $q(Z)$, and $q(\theta)$. When the partial derivative of the free energy functional is taken with respect to $q(Z)$, as presented in Eq. 5.11, the latent parameters, Z , can be updated by holding the model parameters, θ , fixed and optimizing for Z . Likewise, when the partial derivative of the free energy functional is taken with respect to $q(\theta)$, as presented in Eq. 5.12, the model parameters, θ , can be updated by holding the latent parameters, Z , fixed and optimizing for θ . Within Eq. 5.11 and Eq. 5.12 the terms C_z and C_θ are normalizing

⁹For a thorough review on the lower bound maximization for mixture model fitting, the reader is referred to [99].

constants for $q(Z)$ and $q(\theta)$, respectively.

$$q(z)_{t+1} = C_z \int q(\theta)_t \log p(R, Z | \theta) d\theta \quad (5.11)$$

$$q(\theta)_{t+1} = C_\theta p(\theta) \int q(Z)_{t+1} \log p(R, Z | \theta) dZ \quad (5.12)$$

When the variational inference framework, as summarized in Eq. 5.11 and Eq. 5.12, is provided with a data set, R , which conforms to the models discussed in section 5.2.1, the output is a **GMM** which characterizes the underlying probability distribution. Within this chapter, the **GMM** will be utilized to provide an adaptive characterization of the measurement uncertainty model.

5.2.3 ALGORITHM OVERVIEW

Multiple robust state estimation frameworks have been developed, as discussed in Chapter 3. However, it was noted that all of the discussed approaches fall short on at least one front. The primary shortcoming shared by all the approaches is the inability to provide an accurate state estimate when provided with an inaccurate *a priori* measurement error uncertainty model. To confront this concern, a novel **IRLS** formulation is proposed within this section.

The proposed batch covariance estimation (**BCE**) approach is graphically depicted in Fig. 5.2.2, where it is shown that the algorithm is composed of two sections: the initialization, and the robust iteration. To initialize the algorithm, a factor graph is constructed given the *a priori* state and uncertainty information and the set of measurements. Utilizing the constructed factor graph, an initial iteration of the **NLLS** optimizer is used to update the state estimate and calculate a set of measurement residuals.

With the residuals from the initial iteration of the **NLLS** optimizer, the algorithm proceeds into the first robust iteration step. Each robust iteration step commences with the generation of a **GMM** that characterizes the measurement uncertainty through the variational inference framework discussed in section 5.2.2, operating on the most recently calculated set of residuals. With the provided **GMM**, the measurement uncertainty model of the factor graph is updated (i.e., each measurement’s uncertainty model is updated based upon the sufficient statistics of the assigned cluster from the **GMM**). Finally, the state estimate is updated by feeding the modified factor graph into the **NLLS** optimization algorithm. This robust iteration scheme is iterated until a measure of convergence¹⁰ – or the maximum number of iterations¹¹ – has been reached.

Because this framework iteratively updates a **GMM** based measurement uncertainty model, the proposed approach is not only robust to erroneous measurements, but also robust to poor estimates of the *a priori* measurement uncertainty model.

¹⁰To provide a measure of convergence, several criteria could be utilized. For this analysis the change in total error between consecutive iterations was utilized (i.e., the solution has converged to a minimum if the error between consecutive iterations is less than a predefined defined threshold).

¹¹The maximum number of iterations can be selected based upon a number of criteria (e.g., run-time consideration). For the analysis presented within this article, the maximum number of iterations was selected to equal 100.

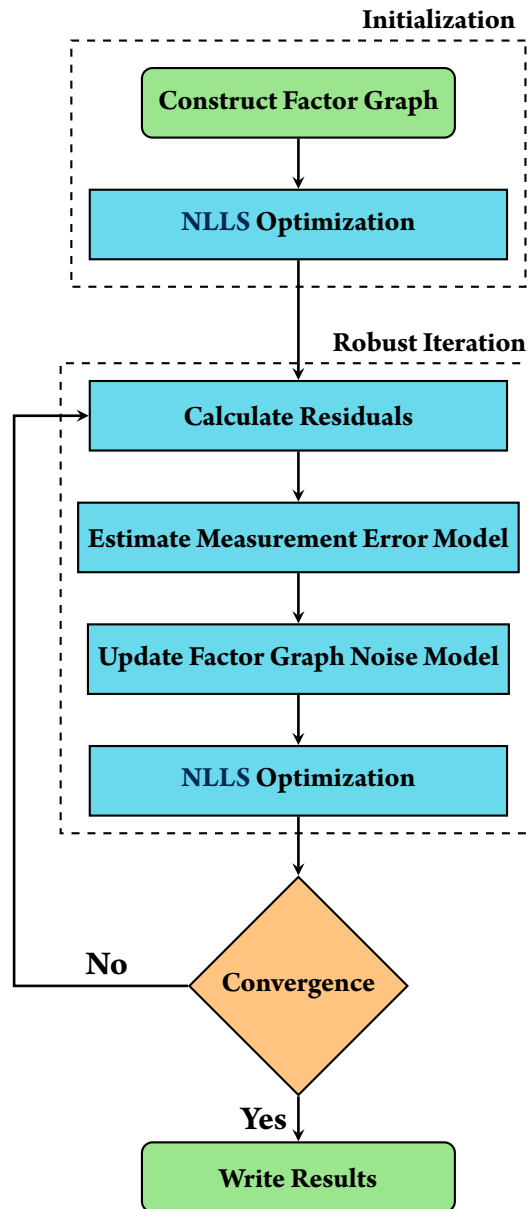


Figure 5.2.2: Schematical overview of the batch covariance estimation algorithm. The proposed approach enables robust estimation through the iterative calculation of the measurement error covariance model, where the measurement error covariance model is characterized by a Gaussian mixture model that is fit to the state estimation residuals.

5.3 EXPERIMENTAL VALIDATION

5.3.1 DATA COLLECTION

To demonstrate the capabilities of the proposed robust optimization framework, an evaluation of the system using global positioning system (GPS) signals from a variety of environments, including open-air, and urban terrain. Due to the varying environments, the measurement uncertainty for each GPS measurements is unknown *a priori* and differs over time. Furthermore, using different qualities of GPS receivers also leads to differing measurement uncertainties.

To test the proposed estimation approach using actual degraded GPS observables, binary in-phase and quadrature (IQ) data in the L₁-band was recorded using a LabSat-3 GPS record and playback device [100] during three kinematic driving tests, as depicted in Fig. 7.3.1, in the Morgantown, WV area. These IQ data were then played back into both a geodetic-grade GNSS receiver (Novatel OEM-638) and an open-source GPS software defined radio (SDR), the GNSS-SDR library [101]. With this experimental setup, the geodetic-grade GNSS receiver was treated as a baseline reference for the best achievable GPS observable quality, and the GPS SDR was used to tune the quality of the observables to two different levels of performance ranging from low-grade to high-grade (i.e,

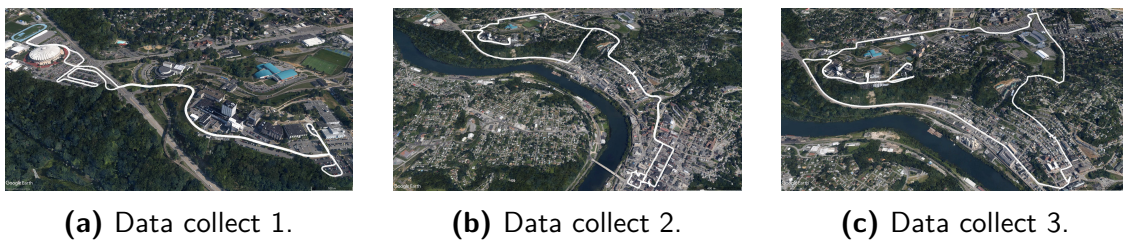


Figure 5.3.1: Ground trace of three kinematic driving data sets collected in the Morgantown, WV area.

close to matching the reference receiver).

By altering the tracking parameters of the GPS SDR, it is possible to vary the accuracy of the GPS pseudorange and carrier-phase observables reported by the receiver. In particular, by changing the tracking loop noise bandwidths and the spacing of the early and late correlators, both the level of thermal noise errors [77] and the susceptibility to multipath errors can be varied [102]. As detailed in [77], the thermal noise of the Phase Lock Loop (PLL) is directly proportional to the square root of the selected noise loop bandwidth, B_ϕ

$$\sigma_{PLL} = \frac{\lambda_{L1}}{2\pi} \sqrt{\frac{B_\phi}{C/N_o} \left(1 + \frac{1}{2TC/N_o}\right)} \quad (5.13)$$

where λ_{L1} is the GPS L1 wavelength, B_ϕ is the carrier loop bandwidth in units of Hz, C/N_o is the carrier-to-noise ratio in Hz, and T is the integration time.

Likewise, the code tracking jitter of the delay lock loop (DLL) is dependent on both the selected noise bandwidth, B_ρ , and the spacing of the early and late correlators, D_{EL} , however, this error model is also dependent on the bandwidth of the radio frequency front-end, B_{fe} , of the receiver and the chiprate of the signal being tracked. That is, if D_{EL} is selected to be at or larger than π multiplied by the ratio of the GPS C/A chipping rate, $R_{C/A}$, and B_{fe} , then the approximation of the code tracking noise jitter (in units of chips) is given as [77]

$$\sigma_{DLL} = \sqrt{\frac{B_\rho}{2C/N_o} D_{EL} \left(1 + \frac{2}{TC/N_o(2 - D_{EL})}\right)} \quad (5.14)$$

if $\left(\frac{R_{C/A}}{B_{fe}} < D_{EL} < \frac{\pi R_{C/A}}{B_{fe}}\right)$, then the code tracking jitter is approximated as [77]

$$\sigma_{DLL} = \left[\frac{B_p}{2C/N_o} \left(\frac{1}{B_{fe}} + \frac{B_{fe} T_c}{\pi - 1} \left(D_{EL} - \frac{1}{B_{fe} T_c} \right)^2 \right) \times \left(1 + \frac{2}{TC/N_o(2 - D_{EL})} \right) \right]^{1/2}.$$

Otherwise, if $\left(D_{EL} < \frac{R_{C/A}}{B_{fe}}\right)$, there is no additional benefit of further reducing D_{EL} , and the code noise error model reduces to a dependence on the front end and noise bandwidths[77].

$$\sigma_{DLL} = \sqrt{\frac{B_p}{2C/N_o} \left(\frac{1}{B_{fe} T_c} \right) \left(1 + \frac{1}{TC/N_o} \right)} \quad (5.15)$$

For the experimental setup used in this study, the LabSat-3 has a B_{fe} equal to 9.66 MHz and the GPS L1 C/A signal has a chiprate of 1.023 MHz, leading to a critical range of $D_{EL} \approx [0.11 \ .33]$. An additional parameter that can be intuitively tuned in the GPS SDR configuration to provide an impact the observable quality is the rate at which the IQ data was sampled, f_s . In this setup, the LabSat-3 has a pre-defined sampling rate of 16.368 Mhz, but the data is down sampled to simulate reduced sampling quality in lower-cost receivers. Using these parameters, we are able to simulate a high-grade and low-grade quality GPS receiver using the parameters shown in Table 5.3.1.

To quantify the accuracy of the generated observations, the zero-baseline test [61] was implemented. A zero-baseline test, as its name suggests, performs a doubled-differenced differential GPS baseline estimation between two sets of GPS receiver data that are known to have an a baseline of exactly zero. That is, in this set-up, the non-zero estimated baseline magnitude is known to be due to discrepancies in the GPS observables reported by the two different receivers. Using the high quality geodetic-grade GPS receiver observables as a reference, the zero-baseline tests provides a

Table 5.3.1: GPS SDR observation tracking configurations utilized within this study.

GPS Quality	f_s (MHz)	D_{EL} (chips)	B_ρ (Hz)	B_ϕ (Hz)
Low	4.092	0.5	2	50
High	16.368	0.2	1	25

Table 5.3.2: Zero-baseline observation comparison for kinematic data sets – see Fig. 7.3.1 – with varied GPS SDR tracking configurations, as specified in Table 5.3.1.

Zero-Baseline Comparison Mean 3D Error (m.)	Collect 1	Collect 2	Collect 3
Low	16.20	37.44	29.43
High	0.59	18.03	17.48

metric to assess the quality of the SDR receiver observables.

The zero-baseline test was implemented by first estimating the reference receiver solution and then configuring the RTKLIB DGPS software [88] to estimate a moving baseline double-differenced GPS solution between the observables reported by the reference receiver and the observables reported by the particular GPS SDR tracking configuration. Because the observables are generated from the same RF front end (i.e., the LabSat-3) and connected to the same antenna, the zero-baseline test results should provided zero difference, in an ideal scenario. The zero-baseline test results for the two resulting GPS SDR cases, which are utilized to experimentally validate the proposed approach within this chapter, are shown in Table 5.3.2, where a large discrepancy can be see between the low-grade and high-grade observations.

The raw IQ recordings, GNSS-SDR processing configurations, and resulting observables in the receiver independent exchange format (RINEX) [103], that are used in this study have all been made publicly available at <https://bit.ly/2vybpgA>.

5.3.2 GPS FACTOR GRAPH MODEL

To enable a self-contained review within the current Chapter, a succinct review of the utilized PPP observation model is provided. However, for a more through review, the reader is referred to the discussion provided in Chapter 4.

To utilize the collected GPS data within the factor graph framework, a likelihood factor must be constructed for the observations. To facilitate the construction of such a factor, the measurement model for the GPS L1-band pseudorange and carrier-phase observations will be discussed, as provided in Eq. 5.16 and Eq. 5.17, respectively. Within the pseudorange and carrier-phase observation models, the X_s term is the known satellites location, X_u is the estimated user location, and $\| * \|$ is the l^2 -norm.

$$\begin{aligned} \rho_{L_1}^i = & \|X_s - X_u\| + c(\delta t_u - \delta t_s) + T_{z,d}M_d(e^l) \\ & + I_{L_1} + \delta_{Rel.} + \delta_{P.C.} + \delta_{D.C.B} + \varepsilon_p^j \end{aligned} \quad (5.16)$$

$$\begin{aligned} \phi_{L_1}^i = & \|X_s - X_u\| + c(\delta t_u - \delta t_s) + T_{z,d}M_d(e^l) \\ & - I_{L_1} + \delta_{Rel.} + \delta_{P.C.} + \delta_{W.U.} + \lambda_{IF}N_{IF}^j + \varepsilon_\phi^j \end{aligned} \quad (5.17)$$

Contained within the models for the pseudorange and carrier-phase observations are several mutual terms, which can be decomposed into three categories. The first category is the propagation medium specific terms (i.e., the troposphere delay $T_{z,d}M_d(e^l)$ [104] and the ionospheric delay I_{L_1} [76]). The second category is the GPS receiver specific terms (i.e., the receiver clock bias δt_u , which must be incorporated into the state vector and estimated, and the observation uncertainty models

ϵ). The final category is the GPS satellite specific terms (i.e., the satellite receiver clock bias δt_s , the differential code bias correction $\delta_{D.C.B}$ [77], the relativistic satellite correction δ_{Rel} [79], and the satellite phase center offset correction $\delta_{P.C}$ [80]).

Contained within the carrier-phase observation model (i.e., Eq. 5.17) are two terms that are not present in the pseudorange observation model. The first term is the carrier-phase windup [81], which can easily be modeled to mitigate its effect. The second term is the carrier-phase ambiguity [105], which can not be easily modeled, and thus, must be incorporated into the state vector and estimated.

Utilizing the provided observation model, the GPS factor graph constraint can be constructed. To begin, it is assumed that the GPS observation uncertainty model is a uni-modal Gaussian. With this assumption, the GPS observations can be incorporated into the factor graph formulation using the mahalanobis distance [83, 106], as provided in Eq. 5.18, where y is the observed GPS measurement, \hat{y} is the modeled observation (calculated using Eq. 5.16 or Eq. 5.17 for a pseudorange or carrier-phase observation, respectively), and Λ is the measurement uncertainty model.

$$\psi(i) = \|y - \hat{y}\|_{\Lambda} \quad (5.18)$$

With the GPS observation factor, the desired states can be estimated using any NLLS algorithm [83], as discussed in Chapter 3. For this application, the state vector is defined in Eq. 5.19, where δP is the 3-D user position state, $T_{z,w}$ is a state used to compensate for troposphere modeling errors, δt_u is the receiver clock bias, and N_* are the carrier-phase ambiguity terms for all observed satellites. For a thorough discussion on the resolution of carrier-phase ambiguity terms within the

factor graph formulation, the reader is referred to [106].

$$X = \begin{pmatrix} \delta P \\ T_{z,w} \\ \delta t_u \\ N_1 \\ \vdots \\ N_n \end{pmatrix} \quad (5.19)$$

To implement the provided model, this chapter benefited from several open-source software packages. Specifically, to enable the implementation of the GPS observation modeling, the GPSTk software package [82] is utilized. For the factor graph construction and optimization, the Georgia Tech Smoothing and Mapping (GTSAM) library [107] is leveraged.

5.3.3 RESULTS

Utilizing the three kinematic data sets — as depicted in Fig. 7.3.1 — with varied receiver tracking metrics (i.e., the low quality and high quality receiver metrics, as discussed in Section 5.3.1, the proposed algorithm was tested with three state estimation techniques. The first algorithm used as a baseline comparison is a batch estimation strategy with an l^2 -norm cost function. The second comparison algorithm is a batch estimator with the dynamic covariance scaling [16] robust kernel. This specific algorithm was selected because it is both a closed-form solution to the switch constraints technique [15], and a specific realization of the robust m-estimator. The final state estimation technique used to for this analysis is the max-mixtures [17] approach with a static measurement uncertainty model.

To generate the reference position solution which is utilized to enable a positioning error comparison, the GNSS observables reported by the reference GNSS receiver (i.e., a NovAtel OEM 638 Receiver) were used. Utilizing the baseline reference GPS measurements, the reference truth solution was generated through an iterative filter-smoother framework implemented within the open-source package RTKLIB [88].

LOW QUALITY OBSERVATIONS

To begin an analysis, the discussed robust estimation techniques are evaluated when the algorithms are provided with low quality observations. As a metric to enable state estimation performance comparisons, the horizontal residual-sum-of-squares (RSOS) positioning error¹² is utilized. Utilizing the horizontal RSOS positioning metric, a solution comparison is provided in the form of a box-plot, as depicted in Fig. 5.3.2. From Fig. 5.3.2, it can be seen that the proposed algorithm significantly reduces the median horizontal positioning error when compared to the reference robust estimation techniques on the three collect data sets – see Table 5.3.3 for specific statistics. In addition to the median RSOS positioning error minimization, it is also noted that the proposed approach either outperforms or performs comparably to the three comparison approaches with respect to positioning solution variance and maximum error, as provided in Table 5.3.3.

To depict the adaptive nature of the measurement uncertainty model within the proposed approach, the estimated pseudorange and carrier-phase uncertainty models (for data collect 1 with low-quality observation) as a function of optimization iteration are depicted in Figures 5.3.3a, and 5.3.3b, respectively. From Fig. 5.3.3b a large change in the structure of the assumed carrier-phase

¹²The horizontal RSOS positioning error is utilized in place of the full 3-D RSOS positioning error due to the inability to accurately model the ionospheric delay with single frequency GPS observations, which imposes a constant bias for all estimations on the vertical positioning error.

Table 5.3.3: Horizontal RSOS positioning error results when low quality observations are utilized. The green and red cell entries correspond to the minimum and maximum statistic, respectively.

(a) Horizontal RSOS positioning error results for data collect 1 when low quality – see Table 5.3.1 for receiver configuration – observations are utilized.

(m.)	L_2	DCS	MM	BCE
median	9.84	10.82	9.13	6.70
variance	0.33	0.21	0.37	0.09
max	14.84	16.10	14.11	11.75

(b) Horizontal RSOS positioning error results for data collect 2 when low quality – see Table 5.3.1 for receiver configuration – observations are utilized.

(m.)	L_2	DCS	MM	BCE
median	5.09	5.02	4.11	3.58
variance	613.13	342.92	673.50	393.32
max	127.29	98.05	132.49	103.64

(c) Horizontal RSOS positioning error results for data collect 3 when low quality – see Table 5.3.1 for receiver configuration – observations are utilized.

(m.)	L_2	DCS	MM	BCE
median	10.98	9.88	8.02	7.31
variance	3.06	1.49	3.36	2.55
max	17.26	17.30	15.06	15.35

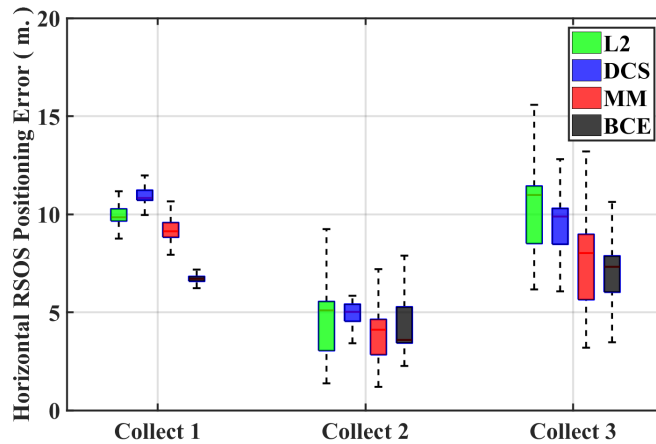


Figure 5.3.2: Box plot of RSOS positioning error for collected kinematic GNSS data sets with low quality – see Table 5.3.1 for GNSS receiver configuration – generated observations. The specific estimator statistics are provided in Table 5.3.3. Within this figure, L_2 is a batch estimator with l^2 cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and BCE is the proposed batch covariance estimation technique.

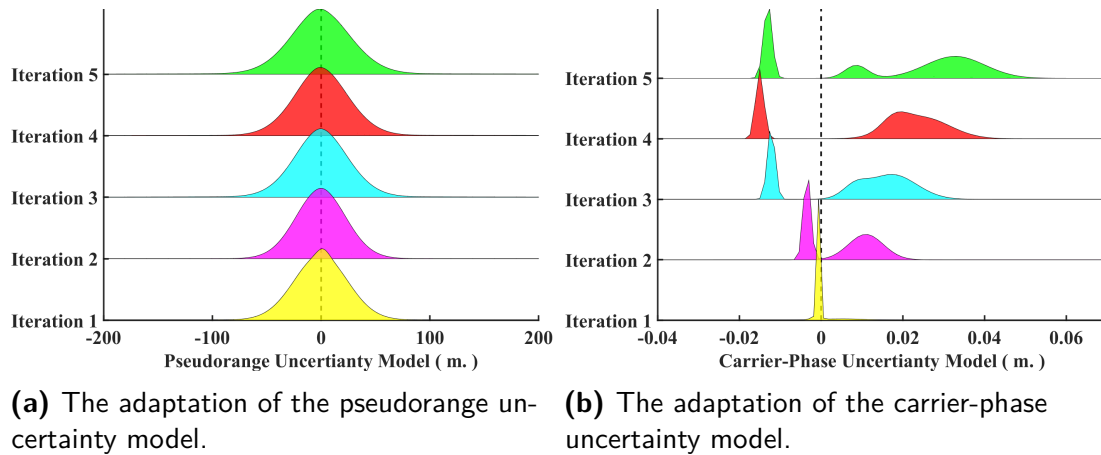


Figure 5.3.3: Evolution of the GNSS observation Gaussian mixture model based uncertainty model for data collect 1 with low-quality observation as a function of optimizer iteration.

measurement uncertainty model can be seen. Specifically, the carrier-phase uncertainty model adapts from the assumed model (i.e, $y \sim N(0, 2.5cm)$) to a highly multimodal GMM at the final iteration, as depicted in Fig. 5.3.3b.

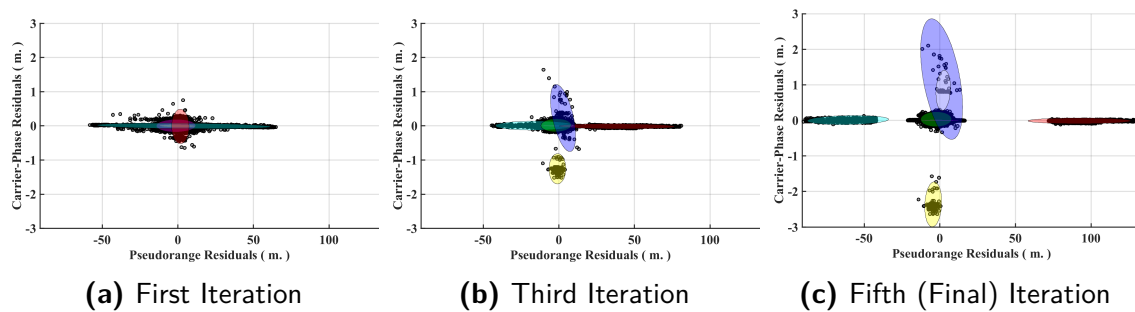


Figure 5.3.4: Progression of GNSS observation residuals with multimodal measurement error covariance – as estimated by the proposed approach which utilizes variational inference to generate a Gaussian mixture model uncertainty representation from the estimators residuals – for data collect 1 with low-quality observations.

HIGH QUALITY OBSERVATION

The high quality observations were also evaluated with the same four state estimation techniques. The horizontal RSOS positioning errors are depicted in box plot format in Figure 5.3.5. From Fig. 5.3.5 it is shown that all four estimators provide similar horizontal positioning accuracy – see Table 5.3.4 for complete statistics.

The comparable horizontal RSOS positioning performance of all four estimation on the provided high quality observations is to be expected. This comparability, when provided with high quality observations, is because the *a priori* assumed model closely resembles the observed model. This means that no benefit is granted by having a robust or adaptive state estimation framework in place.

To verify that the assumed measurement model closely resembles the observed model, the estimated covariance of the proposed approach can be examined for a specific data set (i.e., data collect 1, as depicted in Fig. 5.3.6). For data collect 1 with high quality observations, the measurement uncertainty model estimated by the proposed approach has two modes, as depicted in Fig. 5.3.6. One of the modes — specifically, the mode that characterizes approximately 90% of the measurements — closely resembles the assumed *a priori* measurement uncertainty model. For comparison, the specific values of the *a priori* measurement uncertainty model (Λ), and the estimated measurement uncertainty model ($\hat{\Lambda}$) as provided by the proposed approach, are provided below.

$$\Lambda = \begin{bmatrix} 2.5^2 & 0.0 \\ 0.0 & 0.025^2 \end{bmatrix} \quad \hat{\Lambda} = \begin{bmatrix} 1.7^2 & 0.0 \\ 0.0 & 0.0017^2 \end{bmatrix}$$

Table 5.3.4: Horizontal RSOS positioning error results when high quality observations are utilized. The green and red cell entries correspond to the minimum and maximum statistic, respectively.

(a) Horizontal RSOS positioning error results for data collect 1 when high quality – see Table 5.3.1 for receiver configuration – observations are utilized.

(m.)	L_2	DCS	MM	BCE
median	0.68	0.67	0.75	0.69
variance	0.17	0.17	0.17	0.15
max	6.30	6.30	6.37	6.34

(b) Horizontal RSOS positioning error results for data collect 2 when high quality – see Table 5.3.1 for receiver configuration – observations are utilized.

(m.)	L_2	DCS	MM	BCE
median	0.65	0.65	0.81	0.42
variance	11.27	11.27	11.94	16.60
max	18.29	18.30	18.91	21.73

(c) Horizontal RSOS positioning error results for data collect 3 when high quality – see Table 5.3.1 for receiver configuration – observations are utilized.

(m.)	L_2	DCS	MM	BCE
median	1.44	1.43	1.20	1.24
variance	0.30	0.30	0.30	0.26
max	9.52	17.30	9.43	9.44

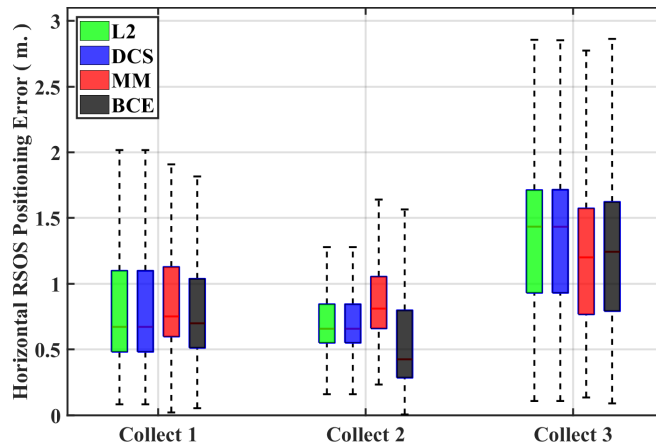


Figure 5.3.5: Box plot of RSOS positioning error for collected kinematic GNSS data sets with high quality – see Table 5.3.1 for GNSS receiver configuration – generated observations. The specific estimator statistics are provided in Table 5.3.4. Within this figure, L_2 , is a batch estimator with L_2 cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and BCE is the proposed batch covariance estimation technique.

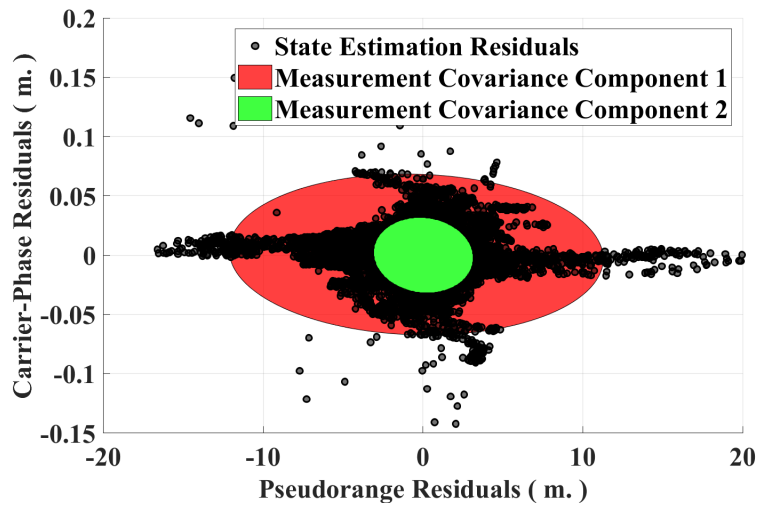


Figure 5.3.6: Measurement residuals with associated BCE estimated measurement error covariance – each ellipse represents 95% confidence – for data collect 1 with high quality observations.

5.3.4 *A PRIORI* INFORMATION SENSITIVITY COMPARISON

To continue the analysis, an evaluation of each estimation framework’s sensitivity to the *a priori* information is provided. Specifically, it is of interest to evaluate the sensitivity of the estimation frameworks to the *a priori* measurement error covariance model.

To enable this estimator sensitivity evaluation, the low-quality observations generated from data collect 1 are utilized. With these generated observations, the estimation framework’s sensitivity is quantified by evaluating the response – which is quantified by the horizontal RSOS positioning error – of each estimator as a function of the *a priori* covariance model. For this study, the *a priori* covariance model will be a scaled version of the assumed measurement error covariance model, as depicted below.

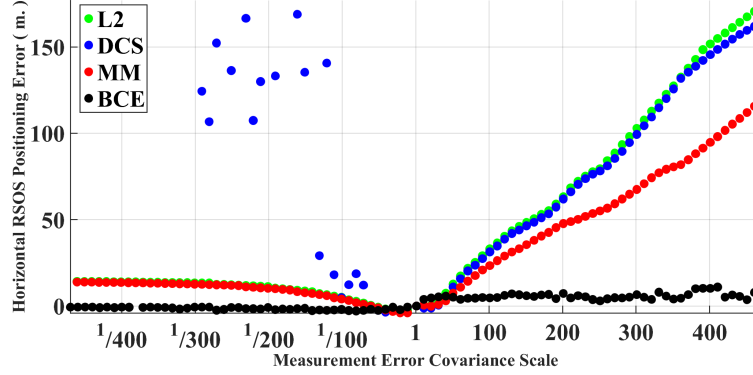


Figure 5.3.7: Sensitivity of the estimation frameworks to the selection of an *a priori* measurement error covariance model. Within this figure, L_2 is a batch estimator with l^2 cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and BCE is the proposed batch covariance estimation technique.

$$\Lambda = s \cdot \begin{bmatrix} 2.5^2 & 0.0 \\ 0.0 & 0.025^2 \end{bmatrix}$$

This sensitivity evaluation is presented graphically within Fig. 5.3.7. From this figure, it is shown that all the estimation frameworks show at least a slight sensitivity to the *a priori* measurement error covariance model; however, some estimation frameworks are significantly more sensitive to perturbation in the *a priori* covariance model than other (e.g., the L_2 and DCS estimators showing the greatest sensitivity). From this figure, it is also shown that the proposed BCE approach is significantly less sensitive to perturbations of the *a priori* measurement error covariance model when compared to the other estimators.

5.3.5 RUN-TIME COMPARISON

To conclude the evaluation of the proposed estimation framework, a run-time comparison is provided. To quantify a run-time evaluation, the wall-clock time (i.e., the total execution time of the utilized estimation framework) divided by the cardinality of the utilized data collect is employed as the metric of comparison. This metric is selected because it enables a run-time comparison regardless of the data collect (i.e., each data collect has a different number of observations so, dividing by the cardinality enables a fair run-time comparison across sets). To implement the run-time comparison, a quad-core Intel i5-6400 central processing unit (CPU) that has a base processing frequency of 2.7 GHz was utilized.

Utilizing the specified hardware and evaluation metric, the run-time comparison is presented within Fig. 7.3.3 for the four estimation algorithms. From Fig. 7.3.3, it is shown that the traditional l^2 approach provides the fastest run-time, with the max-mixtures approach providing comparable results. Additionally, it is shown that the batch covariance estimation approach is the most computationally expensive approach.

The increased computational complexity of the proposed approach is primarily due to the GMM fitting procedure. To reduce the computation requirement of the proposed estimation framework, several research directions could be explored. For example, rather than utilizing all calculated residuals to characterize the measurement uncertainty model, a sub-sampling based approach per iteration could be utilized [108]. This modification has the potential to significantly decrease the run-time of the employed variational clustering algorithm which, in turn, would decrease the run-time of the associated BCE approach.

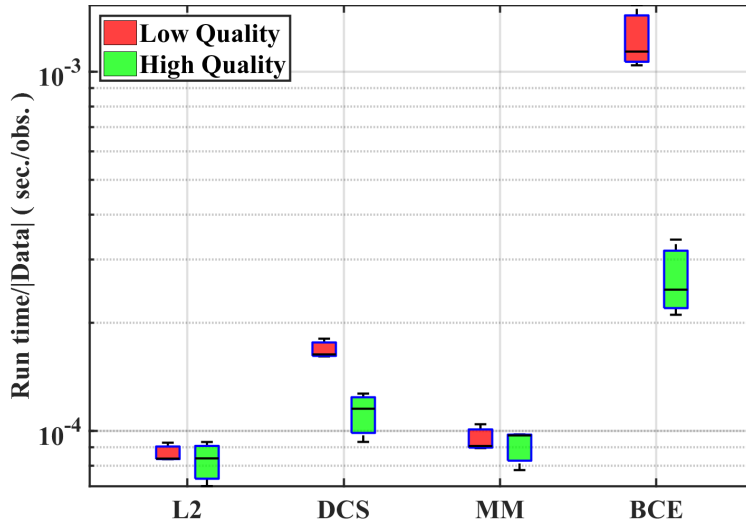


Figure 5.3.8: Wall-clock time per utilized observation for each of the estimation frameworks, where L_2 is a batch estimator with P cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and BCE is the proposed batch covariance estimation technique.

5.4 KEY TAKEAWAY

Several robust state estimation frameworks have been proposed over the previous decades. Underpinning all of these robust frameworks is one dubious assumption. Specifically, the assumption that an accurate *a priori* measurement uncertainty model can be provided. As systems become more autonomous, this assumption becomes less valid (i.e., as systems start operating in novel environments, there is no guarantee that the assumed *a priori* measurement uncertainty model characterizes the sensors current observation uncertainty).

In an attempt to relax this assumption, a novel robust state estimation framework is proposed. The proposed framework enables robust state estimation through the iterative adaptation of the measurement uncertainty model. The adaptation of the measurement uncertainty model is granted through non-parametric clustering of the estimators residuals, which enables the characterization

of the measurement uncertainty via a Gaussian mixture model. This Gaussian mixture model based measurement uncertainty characterization can be incorporated into any non-linear least square optimization routine by only using the single assigned — the assignment of each observation to a single mode within the mixture model is provided by the utilized non-parametric clustering algorithm — component's sufficient statistics from within the mixture model to update the uncertainty model for all observation (i.e., every observations uncertainty model is approximately characterized by the single assigned Gaussian component from within the mixture model).

To verify the proposed algorithm, several GNSS data sets were collected. The collected data sets provide varying levels of observation degradation to enable to characterization of the proposed algorithm on a diverse data set. Utilizing these data sets, it is shown that the proposed technique exhibits improved state estimation accuracy when compared to other robust estimation techniques when confronted with degraded data quality.

6

BCE in an Augmented Data Space

(BCE-AD)

6.1 CHAPTER ABSTRACT

The requirement to generate robust robotic platforms is a critical enabling step to allow such platforms to permeate safety-critical applications (i.e., the localization of autonomous platforms in urban environments). One of the primary components of such a robotic platform is the state estimation engine, which enables the platform to reason about itself and the environment based upon sensor readings. When such sensor readings are degraded traditional state estimation approaches are known to breakdown. To overcome this issue, several robust state estimation frameworks have been proposed. One such method is the batch covariance estimation (BCE) framework. The BCE approach enables robust state estimation by iteratively updating the measurement error uncertainty model through the fitting of a Gaussian mixture model (GMM) to the measurement residuals. This chapter extends upon the BCE approach by arguing that the uncertainty estimation process should be augmented to include metadata (e.g., the signal strength of the associated GNSS observation). The modification of the uncertainty estimation process to an augmented data space is significant because it increases the likelihood of a unique partitioning in the measurement residual domain and thus provides the ability to more accurately characterize the measurement uncertainty model. The proposed batch covariance estimation over an augmented data-space (BCE-AD) is experimentally validated on collected data where it is shown that a significant increase in state estimation accuracy can be granted compared to previously proposed robust estimation techniques.

All developed software and data utilized within this chapter is publicly available at https://github.com/wvunavLab/BCE_AD.

6.2 BCE-AD ALGORITHM FORMULATION

The work presented within this chapter provides an extension to the BCE approach, as detailed in Chapter 5. Specifically, this work removes the assumption that the measurement uncertainty model can be accurately characterized exclusively utilizing information contained within the measurement residuals. Instead, this work argues that the uncertainty estimation problem can effectively be augmented to incorporate metadata, leading to a more accurate characterization of the observation uncertainty model. As shown in Section 6.3 the increase in positioning accuracy achieved by including metadata in the residual classification procedure can be significant.

6.2.1 THE AUGMENTED DATA SPACE

The estimation framework proposed in this chapter is an extension of the BCE methodology as detailed within [46]. In this section there are two primary objectives. Firstly, to expound the differentiating factors between the two implementations. Secondly, to provide an overview of the proposed framework entitled batch covariance estimation over an augmented data-space (BCE-AD).

6.2.2 THE DATA MODEL

To initiate a discussion of the proposed estimation framework, the assumed data model is discussed. In Chapter 5, it is assumed that a given set of residuals $R = \{r_1, r_2, \dots, r_n\}$ with $r_n = y_n - h_n(X) \in \mathbb{R}^d$ can be accurately partitioned into groupings of similar instances. To depict this visually, Fig. 6.2.1a provides the GNSS measurement residuals (i.e., the pseudorange and carrier-phase residuals) for a typical localization application. As illustrated, there is no obvious partitioning of the set.

To increase the likelihood of a unique partitioning, the residual dataset can be projected into a higher-dimensional space [109]. The specific projection utilized within this study is the augmentation of the original dataset with a set of metadata (i.e., for each calculated residual, there is an additional set of features $F = \{f_1, f_2, \dots, f_n\}$ with $f_n \in \mathbb{R}^f$). Where a feature considered for inclusion in the metadata-set is an observed quantity that is known to correlate to the quality of the collected sensor observation¹. To depict the benefit granted by the incorporation of metadata, Fig. 6.2.1b provides the carrier-phase residuals augmented by two additional features (i.e., elevation angle and signal strength of the collected observation). From this figure, it can be seen that the inclusion of metadata provides a more obvious partitioning.

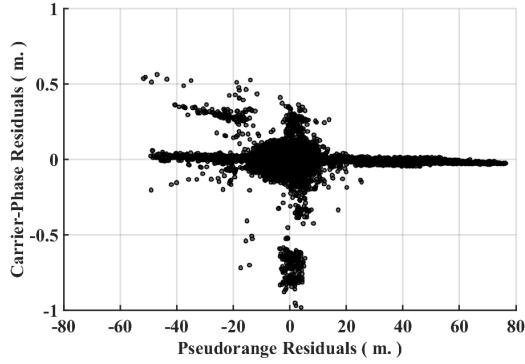
Given this set of augmented data $D = \{d_1, d_2, \dots, d_n\}$ with $d_n = \{r_n, f_n\} \in \mathbb{R}^{d+f}$. It is assumed that the augmented set can be partitioned into similar groupings (i.e., $\bigcup_{m=1}^M C_m = D$), where each group, C_m , can be characterized by a Gaussian distribution (i.e., $C_m \sim N(\mu_m, \Lambda_m)$). With this assumed model, the augmented dataset is fully characterized as a GMM, as depicted in Eq. 6.1

$$D \sim \sum_{m=1}^M w_m N(D | \theta_m) \quad \text{s.t.} \quad \theta_m \triangleq \{\mu_m, \Lambda_m\}, \quad (6.1)$$

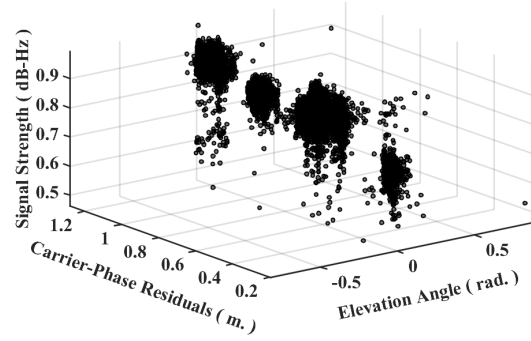
where, m is the number of components in the mixture model², w is the set of mixture weights with the constraint that $\sum_m w_m = 1$, and θ_m is the mixture components sufficient statistics.

¹For GNSS applications, signal strength and elevation angle of a GNSS observation are commonly utilized. Other sensors have similarly useful features (e.g., the mean illumination of an image for computer vision applications)

²The number of mixture components M in the estimated model is not assumed to be known *a priori*. Instead, a truncation level M^* is set [110] (i.e., a maximum number of components, $M^* \geq M$) and the number of utilized components is autonomously selected based upon the variational free energy of the estimated model [94].



(a) Measurement residual data domain (i.e., pseudorange and carrier-phase residuals) extracted after the initial iteration of optimization.



(b) Augmented data domain (i.e., carrier-phase residuals, elevation angle, and signal strength) extracted after the initial iteration of optimization.

Figure 6.2.1: Ground trace for the three kinematic GNSS data collects. All three data collected were recorded in Morgantown, Wv

6.2.3 VARIATIONAL CLUSTERING

As discussed within [46], to fit the a **GMM** to a provided dataset there are two broad classes of algorithms. There are the sampling based frameworks (i.e., Monte Carlo Based approaches [96]), and the optimization based frameworks (i.e., the variational based approaches [59]). Within this study, the variational clustering framework [59, 94] is utilized to reduce the computational complexity of the fitting process.

Specifically, the variational clustering approach is utilized to estimate the model parameters that maximize the log marginal likelihood of the data, as provided in Eq. 7.3

$$\log p(D) = \log \int p(D, \theta, Z) dZ d\theta, \quad (6.2)$$

where $Z = \{z_1, z_2, \dots, z_m\}$ with $z_m \in \mathbb{R}^M$ is a set of assignment variables (i.e., Z provides the explicit assignment of each data instance to a component within the **GMM**). To enable the tractable

computation of the **GMM**, the mean-field assumption [97] (i.e., $p(D, \theta, Z) \approx q(\theta) q(Z)$), is utilized to construct a lower-bound on the true log marginal likelihood.

With the mean-field assumption in place, the **GMM** parameters can be estimated with the variational framework in an iterative fashion, where the sequence of iteration is presented in Eqs. 6.3 and 6.4. The assignment parameters Z are updated by optimizing Eq. 6.3, where C_z is the normalizing constant to $q(z)$. Then model parameters θ are updated by holding the assignment parameters fixed and optimizing Eq. 6.4, where C_θ is the normalizing constant for $q(\theta)$. This iterative process is continued until the Kullback-Leibler divergence between the approximating and true distributions is minimized.

$$q(Z)_{t+1} = C_z \int q(\theta)_t \log p(D, Z | \theta) d\theta \quad (6.3)$$

$$q(\theta)_{t+1} = C_\theta p(\theta) \int q(Z)_{t+1} \log p(D, Z | \theta) dZ \quad (6.4)$$

6.2.4 FEATURE SELECTION

The inclusion of additional features in the assumed data model have the benefit of increasing the likelihood of a unique partitioning in the measurement domain. However, clustering over an augmented dataset (i.e., a dataset with increased dimensionality) could have the detrimental side effect of increased computation complexity. Thus, to remain computationally tractable, the set of augmenting data must be intelligently selected.

The enabling frameworks for intelligent feature selection can be classified as either offline, or online. The offline approach can either be implemented algorithmically or through the utilization of an area expert (i.e., someone with extensive experience with the utilized sensing modality). A

primary drawback of the offline feature selection implementation is a static feature space model (i.e., the utilized features are provided *a priori* and remain fixed over the duration of the estimation process).

Due to the undesirable properties of an offline feature selection approach, an online framework is utilized within this study. Specifically the framework developed within [111] is utilized to autonomously select the most relevant features between each iteration of optimization. In brief, the utilized feature selection algorithm can be described through three primarily steps: 1) construct a nearest neighbor (NN) graph from the provided augmented dataset, 2) conduct eigenvalue decomposition on the NN graph to measure the importance of each feature for partitioning the dataset, 3) perform least angle regression (LARS) [112] on the calculated eigenvectors, to find the most important features. For a more thorough discussion on the feature selection algorithm, the reader is referred to [111].

6.2.5 ALGORITHM OVERVIEW

With the previously detailed topics, the discussion can now proceed to an overview of the proposed estimation framework. From Fig. 6.2.2, it is shown that the proposed algorithm is comprised of two primary segments. The first component is the initialization of the estimator. This process begins by constructing the factor graph representation of the NLLS optimization problem – as thoroughly detailed within [27] – from the *a priori* state and uncertainty information and the provided observations. Then, an initial iteration of optimization is conducted with a NLLS estimation algorithm (e.g., for this study, the Levenberg-Marquardt [113] implementation was utilized) to update the *a priori* state estimate.

The second component of the proposed framework commences with the calculation of the mea-

surement residuals given the previously estimated set of states. The set of calculated residuals are augmented with the provided set of metadata, as discussed in section 6.2.2. Using the augmented dataset for the current iteration of optimization, the most relevant features are selected using the framework discussed in section 6.2.4. Using the relevant set of features, the variational cluster framework – as discussed in section 6.2.3 – is utilized to assign each instance in the augmented dataset to a component within an estimated **GMM**.

Utilizing the partitioning of the augmented data space, the measurement error uncertainty model can be calculated. This is achieved by first partitioning the measurement residuals with the assignment vector Z estimated in the augmented data space. Then, the sufficient statistics for each grouping in the measurement residual domain are calculated. With this m -dimensional **GMM** in the measurement residual domain, the measurement uncertainty model of the factor graph is updated (i.e., each measurement's uncertainty model is updated to the sufficient statistics of the assigned **GMM** component). With the updated measurement uncertainty model, a new iteration of optimization is conducted. This process is iterated until a measure of convergence (e.g., the error decrease between consecutive iterations is less than a user defined threshold) – or a limit on the number of iterations – has been reached.

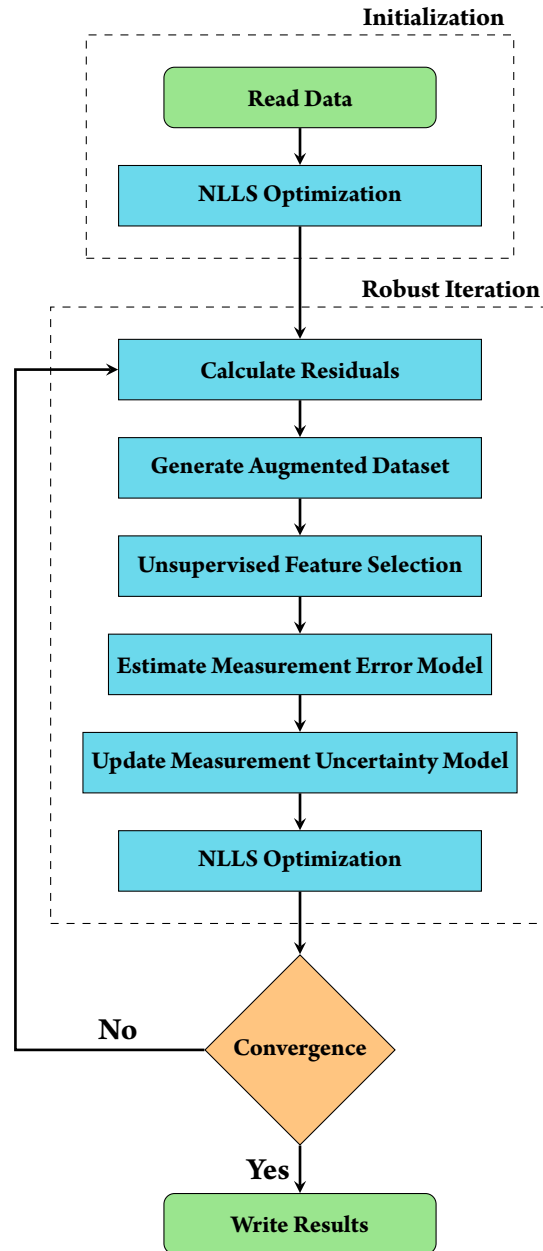


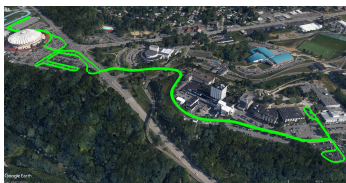
Figure 6.2.2: Overview of the proposed robust optimized algorithm. The proposed approach enables robust state estimation through the iterative estimation of the – possible multimodal – measurement error covariance model, where the measurement error covariance model is estimated by clustering over an augmented data space constructed from the previous iteration of optimization.

6.3 EXPERIMENTAL VALIDATION

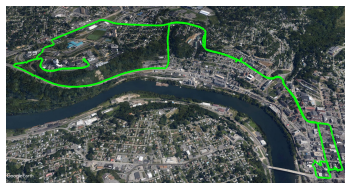
6.3.1 DATA COLLECTION

To enable the validation of the proposed estimation framework, the three kinematic GNSS data collects, as previously discussed in Chapter 5, are utilized. A discussion of the collection process is briefly described below for the sake of completeness; however, the reader is referred to Section 5.3.1 for a more thorough review. The ground-trace for the three driving data collects is provided in Fig. 6.3.1. For each of the depicted data collects, the GNSS binary IQ observations were recorded with a LabSat-3 GPS record and playback device [100]. The collected IQ data were then played back into two GNSS receivers: a geodetic-grade (Novatel OEM-638) GNSS receiver, and an open-source GPS SDR.

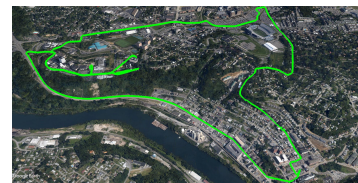
The observations generated by the geodetic-grade receiver, in conjunction with the observations generated by an additional static GNSS receiver, were utilized to calculate a reference positioning solution for each data collect. Where, the reference solution was calculated through a differential (i.e., real time kinematic (RTK)) filter-smoother framework implemented within the open-source package RTKLIB [88].



(a) Ground trace for data collect 1.



(b) Ground trace for data collect 2.



(c) Ground trace for data collect 3.

Figure 6.3.1: Ground trace for the three kinematic GNSS data collects. All three data collected were recorded in Morgantown, Wv

The observations generated by the GPS SDR are utilized by the estimation framework to validate the proposed methodology. These observations are utilized for validation because they were intentionally degraded by altering the GNSS tracking parameters. For a more detailed discussion on the specific GNSS tracking parameter settings, the reader is referred to Section IV of [46].

With these three datasets, the proposed approach (i.e., BCE-AD) is validated against four additional estimators. The first comparison algorithm is a traditional (i.e., non-robust) estimation framework that utilizes a l^2 -norm cost function. Additionally, the DCS robust estimation framework [16] is utilized. The third comparison algorithm is the MM estimation framework [17] with a static measurement uncertainty model. The final comparison algorithm is the BCE approach [46].

6.3.2 RESULTS

To begin the evaluation of the proposed methodology, the horizontal residual-sum-of-squares (RSOS) positioning error for the three data collects will be examined. This evaluation is presented visually in Fig. 6.3.2, in the form of a box-plot. From Fig. 6.3.2 it is shown that the proposed BCE-AD significantly reduces the median horizontal RSOS positioning error for data collects 1 and 3, and performs comparably well to the other robust estimators on data collect 2. The specific statistics representing the analysis presented in Fig. 6.3.2 are provided in Table 6.3.1.

To examine the reason why the BCE-AD approach provided no additional positioning performance benefit when compared to the BCE approach for data collect 2, the measurement domain partitioning for the two approaches can be evaluated. This evaluation is provided in Fig. 6.3.3 for all three data collects. From the provided figure, it can be hypothesized that no additional positioning accuracy is granted when the measurement residuals can be easily partitioned utilizing exclusively the information in the measurement residual domain. This is depicted visually by evaluating

Table 6.3.1: Horizontal RSOS positioning error results when low quality observations are utilized. The green and red cell entries correspond to the minimum and maximum statistic, respectively.

(a) Horizontal RSOS positioning error results for data collect 1 when low quality – see Table 5.3.1 for receiver configuration – observations are utilized.

(m.)	L_2	DCS	MM	BCE	BCE-AD
median	9.84	10.82	9.13	6.70	2.14
variance	0.33	0.21	0.37	0.09	0.16
max	14.84	16.10	14.11	11.75	7.79

(b) Horizontal RSOS positioning error results for data collect 2 when low quality – see Table 5.3.1 for receiver configuration – observations are utilized.

(m.)	L_2	DCS	MM	BCE	BCE-AD
median	5.09	5.02	4.11	3.58	3.58
variance	613.13	342.92	673.50	393.32	393.32
max	127.29	98.05	132.49	103.64	103.64

(c) Horizontal RSOS positioning error results for data collect 3 when low quality – see Table 5.3.1 for receiver configuration – observations are utilized.

(m.)	L_2	DCS	MM	BCE	BCE-AD
median	10.98	9.88	8.02	7.31	3.25
variance	3.06	1.49	3.36	2.55	3.25
max	17.26	17.30	15.06	15.35	12.64

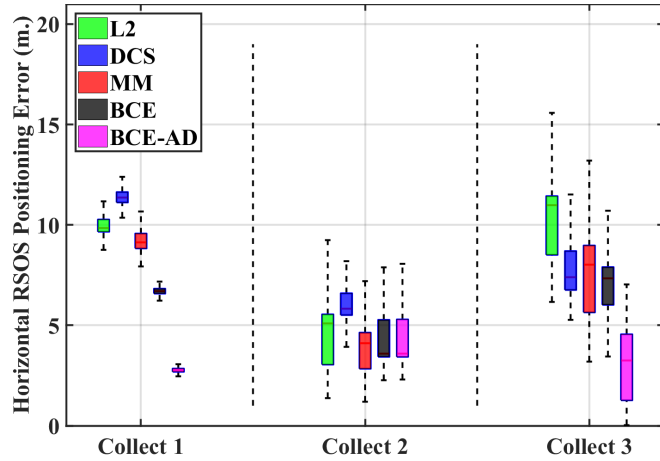


Figure 6.3.2: Horizontal RSOS positioning error for collected GNSS data sets. The specific estimator statistics are provided in Table 6.3.1. Within this figure, L_2 , is a batch estimator with L_2 cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, BCE is the batch covariance estimation technique, and BCE-AD is the proposed batch covariance estimation technique over an augmented data space.

Fig. 6.3.3c, where it is shown that the clustering utilizing only the information in the measurement domain can easily partition the provided measurement residuals. However, when evaluating the measurement residuals for data collects 1 and 3 (i.e., the datasets where the greatest positioning performance benefit is granted by the BCE-AD approach), it is seen that clustering over the two different domains provides significantly different partitionings.

To conclude the evaluation of the proposed approach, the selected features for each dataset are evaluated. This evaluation is depicted visually in Table 6.3.2. From this visual, it can be noted that the most relevant features not only vary from dataset to dataset, but also from one optimization iteration to the next within a given dataset. Additionally, the hypothesis presented in the previous paragraph (i.e., that the BCE-AD only grants increased positioning performance when sufficient information is not present in the measurement residual domain to partition the residuals) is further verified in Table 6.3.2b where it is shown that only features utilized for data collect 2 are the measurement residuals.

(a) Utilized features per iteration of optimization for data collect 1.

	Iter 0.	Iter 1.	Iter 2.	Iter 3.	Iter 4.	Iter 5.	Iter 6.	Iter 7.	Iter 8.
ρ	Green	Red	Green	Green	Green	Green	Green	Green	Green
Φ	Green	Green	Green	Green	Green	Green	Green	Green	Green
El.	Red	Green	Red	Red	Green	Green	Green	Green	Green
Az.	Red	Red	Red	Red	Red	Red	Red	Red	Red
SS.	Red	Green	Red	Red	Green	Green	Green	Green	Green

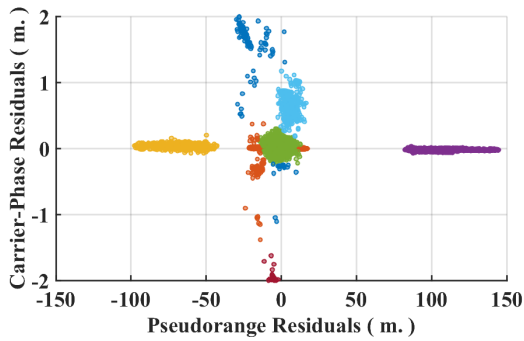
(b) Utilized features per iteration of optimization for data collect 2.

	Iter 0.	Iter 1.	Iter 2.	Iter 3.	Iter 4.	Iter 5.	Iter 6.
ρ	Green	Green	Green	Green	Green	Green	Green
Φ	Green	Green	Green	Green	Green	Green	Green
El.	Red	Red	Red	Red	Red	Red	Red
Az.	Red	Red	Red	Red	Red	Red	Red
SS.	Red	Red	Red	Red	Red	Red	Red

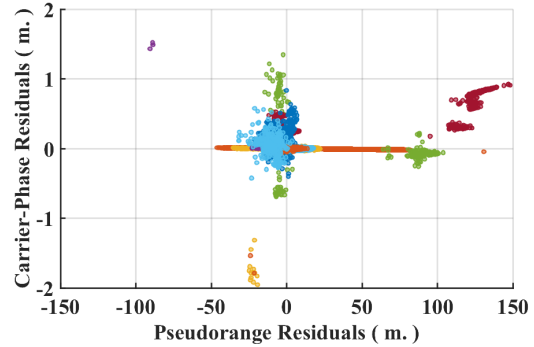
(c) Utilized features per iteration of optimization for data collect 3.

	Iter 0.	Iter 1.	Iter 2.	Iter 3.	Iter 4.
ρ	Green	Green	Green	Green	Green
Φ	Green	Green	Green	Green	Green
El.	Red	Red	Red	Green	Red
Az.	Red	Red	Red	Red	Red
SS.	Red	Green	Green	Green	Green

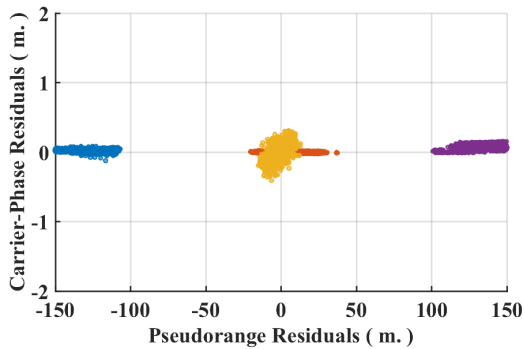
Table 6.3.2: Utilized features for each iteration of optimization with the BCE-AD approach. The green and red cell entries correspond to the utilized and non-utilized features, respectively. Within the table, ρ is the pseudorange residual, Φ is the carrier-phase residual, El. is the elevation angle, Az. is the azimuth angle, and SS. is the signal strength.



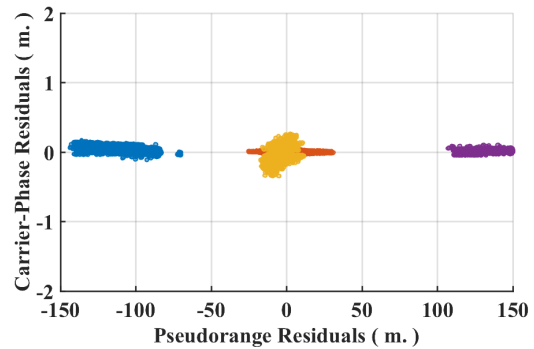
(a) BCE based measurement residual domain partitioning for data collect 1.



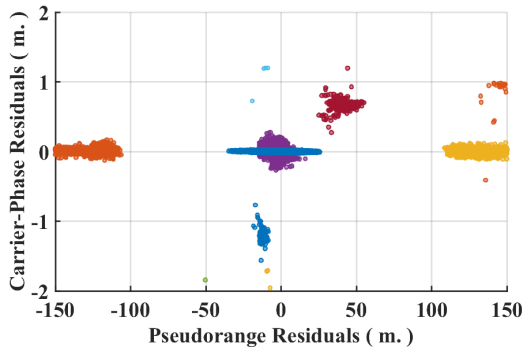
(b) BCE-AD based measurement residual domain partitioning for data collect 1.



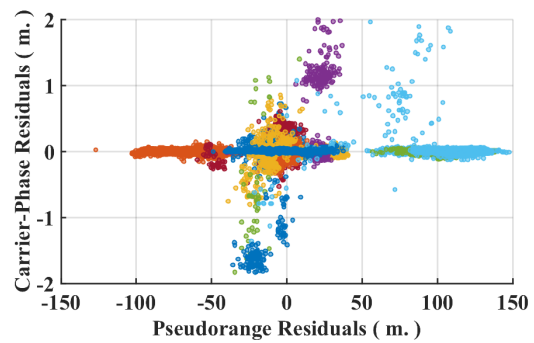
(c) BCE based measurement residual domain partitioning for data collect 2.



(d) BCE-AD based measurement residual domain partitioning for data collect 2.



(e) BCE based measurement residual domain partitioning for data collect 3.



(f) BCE-AD based measurement residual domain partitioning for data collect 3.

Figure 6.3.3: Partitioning of the measurement residual domain with the BCE approach (see Figs. 6.3.3a, 6.3.3c, and 6.3.3e), and the BCE-AD approach (see Figs. 6.3.3b, 6.3.3d, and 6.3.3f), for the three kinematic GNSS data collects at the final iteration of optimization.

6.4 KEY TAKEAWAY

This chapter presents an extension of the previously proposed batch covariance estimation (BCE) technique to enable robust state estimation. The BCE approach enables robust state estimation through the iterative estimation of a measurement error uncertainty model based upon the previous iterations measurement residuals. Where, the estimated measurement error uncertainty model is characterized by a Gaussian mixture model (GMM) which is fit to the measurement residuals through variational clustering. After fitting the GMM to the current optimization iterations residuals, the uncertainty model of each observation is update to the sufficient statistics of the assigned cluster within the GMM.

The approach proposed within this work extends the BCE approach on one front. Specifically, it removes the assumption that the measurement error uncertainty model can be accurately characterized utilizing information exclusively from within the measurement residual domain. Instead, this chapter argues that the uncertainty estimation process should be augmented to include additional signal quality metrics (SQMs). The modification of the uncertainty estimation process to an augmented data space increases the likelihood of a unique partitioning in the measurement residual domain and thus provides the ability to more accurately characterize the measurement uncertainty model.

To verify the proposed batch covariance estimation over and augmented data space (BCE-AD) approach, three GNSS data sets were utilized. The utilized data sets provide varying levels of degradation to quantify the robustness of the proposed algorithm against other state-of-the-art robust estimators. Utilizing these data sets, it is shown that the proposed approach provides comparable or improved state estimation accuracy when compared to other robust estimation techniques.

7

Incremental Covariance Estimation (ICE)

7.1 CHAPTER ABSTRACT

Recent advances in the fields of robotics and automation have spurred significant interest in robust state estimation. To enable robust state estimation, several methodologies have been proposed. One such technique, which has shown promising performance, is the concept of iteratively estimating a Gaussian Mixture Model (GMM), based upon the state estimation residuals, to characterize the measurement uncertainty model. Through this iterative process, the measurement uncertainty model is more accurately characterized, which enables robust state estimation through the appropriate de-weighting of erroneous observations. This approach, however, has traditionally required a batch estimation framework to enable the estimation of the measurement uncertainty model, which is not advantageous to robotic applications. In this chapter, an efficient, incremental extension to the measurement uncertainty model estimation paradigm is proposed. The incremental covariance estimation (ICE) approach, as detailed within this paper, is evaluated on several collected data sets, where it is shown to provide a significant increase in localization accuracy, when compared to other state-of-the-art robust, incremental estimation algorithms.

7.2 THE ICE ALGORITHM FORMULATION

To facilitate a discussion of the proposed incremental covariance estimation (ICE) framework the assumed data model is first explained. Then, a method for incremental measurement uncertainty model adaptation is presented. Finally, pull the previously mentioned topics together, the discus-

All developed software and data utilized within this chapter is publicly available at <https://github.com/wvunavLab/ICE>.

sion concludes with an overview of the proposed ICE framework.

7.2.1 DATA MODEL

As calculated by the estimator, a set of state estimation residuals $\mathbf{R} = \{r_1, r_2, \dots, r_N \mid r_n \triangleq y_n - h_n(X)\}$ is provided. The set of state estimation residuals can be characterized by a GMM, which, for this work, will act as the measurement uncertainty model, GMM_g . As proposed within [17], with the intent to minimize the computation complexity of the optimization problem, the GMM can be reduced to selecting the most likely component from the mixture model to approximately characterize each observation, as depicted in Eq. 7.1 where μ_m is the components mean and Λ_m is the components covariance.

$$r_n \sim \max_m w_m N(r_n \mid \theta_m) \quad \text{s.t.} \quad \theta_m = \{\mu_m, \Lambda_m\} \quad (7.1)$$

For this work, it is additionally assumed that the set of residuals, \mathbf{R} , can be partitioned into two distinct groups. The first group is the set of all residuals which sufficiently adhere to the *a priori* covariance model (i.e., do not deviate sufficiently from the most likely component within GMM_g), which will be indicated by the set \mathbf{R}_I . While, the second group is the set residuals which do not sufficiently adhere to the *a priori* covariance model, which will be indicated by the set \mathbf{R}_O .

To quantify the level of adherence to the *a priori* uncertainty model, the z-test, as provided in Eq. 7.2, is employed. Within Eq. 7.2 μ , and σ are the mean and standard deviation of the most likely component from GMM_g for the state estimation residual r_n . Utilizing the z-test as a metric to quantify the level of agreement between the set of state estimation residual and the *a priori* uncertainty

model, the two groupings can more concretely define as, $\mathbf{R}_I = \{r \mid r \in \mathbf{R}, Z(r, \varphi) < T_r\}^1$ and $\mathbf{R}_O = \{r \mid r \in \mathbf{R}, r \notin \mathbf{R}_I\}$.

$$Z(r_n, \varphi) = \frac{r_n - \mu}{\sigma} \quad \text{s.t.} \quad \varphi \triangleq \{\mu, \sigma\} \quad (7.2)$$

7.2.2 UNCERTAINTY MODEL ADAPTATION

By definition, the set \mathbf{R}_O is not accurately characterized by GMM_g , thus, it is desired to adapt the uncertainty model to more accurately represent the new observations. To enable the adaptation of the uncertainty model, a two step procedure is utilized. This procedure starts by estimating a new GMM , which will be indicated by GMM_n , based solely on the set \mathbf{R}_O . Then, GMM_n is merged into the prior model (i.e., GMM_g) to provide a more accurate characterization the measurement uncertainty model. This procedure is elaborated upon in Section 7.2.2 and Section 7.2.2, respectively.

VARIATIONAL CLUSTERING

To estimate GMM_n , the set of model parameters which maximizes the log marginal likelihood, as depicted in Eq. 7.3, must be calculated. In Eq. 7.3, θ is the set of mean vectors and covariance matrices which define the new GMM , and \mathbf{Z} is an assignment variable (i.e., the variable \mathbf{Z} assigns each $r \in \mathbf{R}_O$ to a specific component within the model).

$$\log p(\mathbf{R}_O) = \log \int p(\mathbf{R}_O, \theta, \mathbf{Z}) d\mathbf{Z} d\theta \quad (7.3)$$

¹ T_r is a user defined parameter that encodes the acceptable amount an observation can deviation from the *a priori* model in terms of multiples of the standard deviation.

In general, the integral presented in Eq. 7.3 is computational intractable [95]. Thus, a method of approximate integration must be implemented. For this work, the variational inference² [94, 95] approach is utilized primarily due this class of algorithms run-time performance when compared to sampling based approaches (i.e., Monte Carlo methods [96]).

EFFICIENT GMM MERGING

To enable the second step of the measurement uncertainty model adaptation (i.e., the merging of GMM_n into the prior model GMM_g), an implementation of the algorithm presented in [114] is utilized. To provide a description of the approach, let's evaluate the equivalence between $g_n \triangleq \{w_n, \mu_n, \Lambda_n\} \in \text{GMM}_n$ (e.g., the first component in GMM_n) and $g_g \triangleq \{w_g, \mu_g, \Lambda_g\} \in \text{GMM}_g$ (e.g., the first component in GMM_g).

To test the equivalence, first the set of observations $\mathbf{R}_{\mathbf{O},g_n} \subseteq \mathbf{R}_{\mathbf{O}}$ that correspond to set of state estimation residuals that are characterized by component g_n are extracted. Utilizing $\mathbf{R}_{\mathbf{O},g_n}$, it is desired to check if the set of state estimation residuals has an equivalent covariance to the hypothesis covariance model (i.e., we want to see if $\Lambda_n = \Lambda_g$, where $\Lambda_n = \text{cov}(\mathbf{R}_{\mathbf{O},g_n})$ and Λ_g is the hypothesis covariance from g_g).

To determine if the two GMM components have an equivalent covariance model, first the set of observations $\mathbf{R}_{\mathbf{O},g_n}$ are transformed with Cholesky decomposition of the hypothesis covariance³. This transformation provides us with a new data set, defined as $\mathbf{Y} = \{y = L^{-1}r \mid r \in \mathbf{R}_{\mathbf{O},g_n}, \Lambda_g = LL^T\}$.

Utilizing the transformed set of state estimation residuals \mathbf{Y} , the W -statistic [115] can be con-

²To enable the implementation of the ICE approach in software, the **libcluster** [94] software library was utilized.

³This whitening process is conducted because the covariance test is only valid for unit covariance matrices.

structured, as provided in Eq. 7.4, to test the equivalence of covariance matrices. Within Eq. 7.4, $\Lambda_y = \text{cov}(\mathbf{Y})$, m is the cardinality of the set \mathbf{Y} (i.e., $m = |\mathbf{Y}|$), and d is the dimension the state estimation residuals (i.e., $y_m \in \mathbb{R}^d$).

$$W = \frac{1}{d} \text{Tr}((\Lambda_y - I)^2) - \frac{d}{m} \left(\frac{1}{d} \text{Tr}(\Lambda_y) \right)^2 + \frac{d}{m} \quad (7.4)$$

The W -statistic is known to have an asymptotic χ^2 distribution with degrees of freedom $d(d + 1)/2$, as depicted in Eq. 7.5. Thus, a Chi-square test with a user defined critical value is utilized to test the equivalence of covariance matrices.

$$\frac{mWd}{2} \sim \chi_{d(d+1)/2}^2 \quad (7.5)$$

To test the equivalence of mean vectors, the T -statistic [116], as provided in Eq. 7.6, is utilized. Within Eq. 7.6, μ_n is the mean of the component of \mathbf{GMM}_n , and μ_g is the mean vector of the component of \mathbf{GMM}_g . The T -statistic is utilized to test the equivalence of mean vectors because it is known to have an asymptotic F distribution, as depicted in Eq. 7.7. Thus, an F-test with user defined critical value is utilized to test the equivalence of mean vectors.

$$T^2 = m \left\| \mu_n - \mu_g \right\|_{\Lambda_y} \quad (7.6)$$

$$\frac{m-d}{d(m-1)} T^2 \sim F_{d,m-d} \quad (7.7)$$

If both the mean and covariance of two components are found to be equivalent, then the new component g_n is merged with the prior component g_g to adapt the measurement uncertainty model

GMM_g . To adapt the measurement uncertainty model, the mean, covariance and weighting can be updated, as presented in Eqs. 7.8, 7.9, and 7.10, respectively. Within Eqs. 7.8, 7.9, and 7.10, N is the total number of points which are characterized by GMM_g , M is the total number of points which are characterized by GMM_n , and m is the number of points which are characterized by component g_n .

$$\mu = \frac{Nw_g\mu_g + m\mu_n}{Nw_g + m} \quad (7.8)$$

$$\Lambda = \frac{Nw_g\Lambda_g + m\Lambda_n}{Nw_g + m} + \frac{Nw_g\mu_g\mu_g^T + m\mu_n\mu_n^T}{Nw_g + m} - \mu\mu^T \quad (7.9)$$

$$w = \frac{Nw_g + m}{N + M} \quad (7.10)$$

If the new component g_n does not match a component within GMM_g , then the mean and covariance of g_n is added to GMM_g . When the new component is added to GMM_g the weighting vector is updating, as presented in Eq. 7.11, where N , M , and m are as defined above. When the new component is added, the weighting for all of the remaining components in GMM_g are updated according to Eq. 7.12.

$$w = \frac{m}{N + M} \quad (7.11)$$

$$w = \frac{Nw_g}{N + M} \quad (7.12)$$

Through the utilization of the mixture model merging approach developed within [114], and

outlined in this section, the measurement uncertainty model can be adapted online. This adaptation is conducted without the need for storing all previous state estimation residuals (i.e, only the most recent residuals \mathbf{R}_O which do not adhere to the *a priori* model are required), which dramatically reduces the computational and memory cost of the proposed approach.

7.2.3 ALGORITHM OVERVIEW

With the discussion provided in the previous sections, the conversation can now turn to an overview of the proposed robust estimation framework. To facilitate a discussion, a graphical overview of the ICE framework is depicted in Fig. 7.2.1.

From Fig. 7.2.1, it is shown that the ICE algorithm starts at each epoch by calculating the set of state estimation residuals \mathbf{R}_t from the current set of observations \mathbf{Y}_t . As discussed within Section 7.2.1, this set of state estimation residuals \mathbf{R}_t can be partitioned into two distinct groups (i.e., the set of state estimation residuals which correspond to erroneous observations $\mathbf{R}_{O,t}$, and the set of state estimation residuals which correspond to observations that adhere to the *a priori* model $\mathbf{R}_{I,t}$) through the utilization of the *z*-test.

With the set $\mathbf{R}_{O,t}$, the previous set of state estimation residuals which correspond to erroneous observations \mathbf{R}_O is appended. If the length of \mathbf{R}_O is greater than a user defined threshold⁴ (i.e., if $|\mathbf{R}_O| > T_c$), the set is utilized to modify the measurement uncertainty model, as described in Section 7.2.2. After the adaptation of the uncertainty model, the set \mathbf{R}_O is cleared and the set of observations which adhere to the *a priori* model $\mathbf{R}_{I,t}$ are incorporated. With the incorporation of the new observations, a new state estimate is provided, following the discussion provided in Section 3.2.

⁴Several factors can affect the specific realization of this threshold (e.g., the expected dynamics of the environment, or the number of observations per epoch).

If the length of set of state estimation residuals, which correspond to erroneous observations, \mathbf{R}_O is less than a user defined threshold, then the uncertainty model is not adapted for the current epoch. Instead, the previous measurement uncertainty model is utilized to incorporate the new set of observations which adhere to the *a priori* model. With the new observations incorporated, a new state estimated is provided, as described in Section 3.2. This process is continued in an iterative

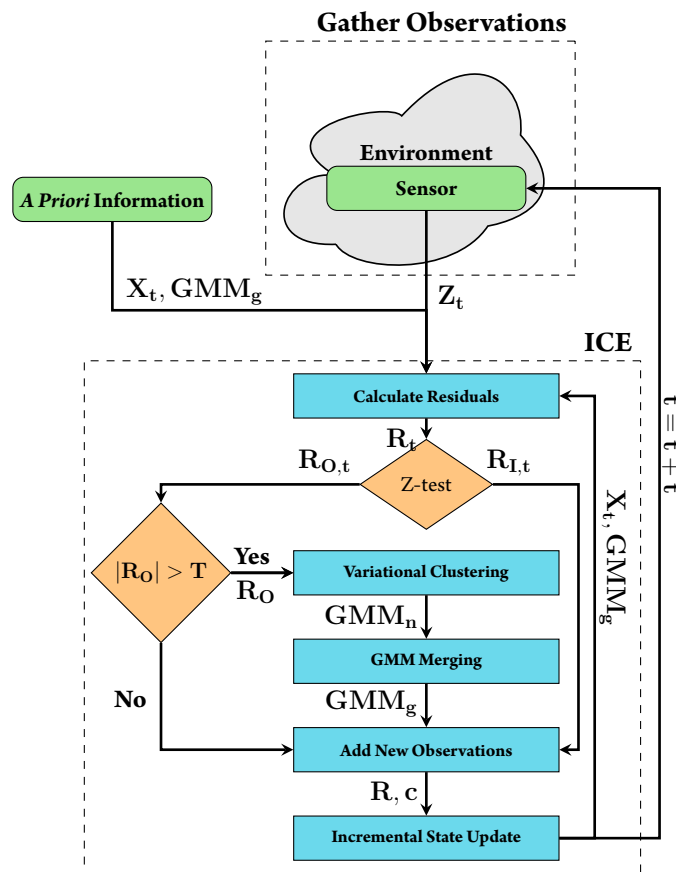


Figure 7.2.1: Graphical depiction of the proposed incremental covariance estimation (ICE) algorithm. The proposed approach enables efficient, incremental, and robust state estimation through the iterative adaptation of the measurement uncertainty model, based upon the state estimation residuals that correspond to erroneous observations.

fashion for as long as needed (e.g., until the data collection terminates).

7.3 RESULTS

7.3.1 DATA COLLECTION

To conduct an evaluation of the proposed robust estimation framework, a collection of three kinematic GNSS data sets, as discussed in Section 5.3.1, is utilized. These GNSS data sets, as can be visualized through their ground traces, which are shown in Fig. 7.3.1, were made publicly available and are described within [46].

For these data collects, the binary IQ data in the L₁-band was recorded. By recording the IQ data in place of the GNSS receiver dependent observations (i.e., the pseudorange and carrier-phase observables), the same data collect can be utilized to generate several sets of observations with varying levels of degradation after playing back through a software defined GNSS receiver [101] with different sets of tracking parameters. Specifically, the receiver dependent observations can be generated off-line by playing the IQ data into a GNSS receiver, where the level of degradation is varied by changing the GNSS receiver's tracking parameters (i.e., changing the bandwidth of the phase lock loop (PLL), the DLL and the correlator spacing). For a detailed discussion on the impact that the GNSS receiver tracking parameters can have on the quality of the generated observables, the reader is referred to [77, 102], which is reviewed in [46].

For this study, two sets of observations are generated (i.e., a low-quality and high-quality data set) for each of the data collects. The specific GNSS receiver tracking parameters utilized to generate the low-quality and high-quality observations are provided within Table 5.3.1.

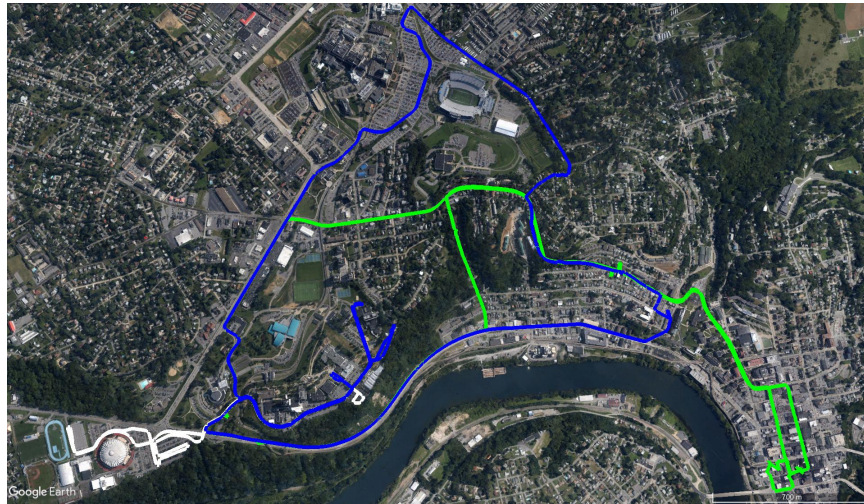


Figure 7.3.1: Ground trace for the three utilized GNSS data sets. The white trace corresponds to data collect 1, the green trace corresponds to data collect 2, and the blue trace corresponds to data collect 3.

7.3.2 EVALUATION

Utilizing these data collects, an evaluation of the proposed methodology can be conducted. To provided a comparison for the proposed approach, three additional estimation frameworks will be utilized. The first comparison methodology is the traditional l^2 -norm based estimator. The second comparison methodology is the *MM* approach, which has a static measurement error covariance model (i.e., a fixed two component measurement error covariance model). The final comparison methodology is the *DCS* approach, where the *DCS* approach is utilized because it is both a closed form version of switchable constraints and a specific implementation of an *m*-estimator [?]. All of the utilized estimators are built upon the *iSAM*₂ algorithm [45], as implemented within the Georgia Tech Smoothing and Mapping (*GTSAM*) library [117].

POSITIONING PERFORMANCE

To start an evaluation, the localization performance of the estimation frameworks will be assessed. To enable the assessment of the localization performance, a reference ground-truth must first be established. To generate this ground-truth, a differential GNSS solution (i.e., RTK⁵) is utilized, which is known to provide centimeter level localization accuracy [77].

With the RTK generated reference ground-truth solution, the localization performance of the four estimation frameworks, when low-quality observations are utilized, is provided in Table 7.3.1⁶. From Table 7.3.1, it can be seen that all three of the robust estimation frameworks provided a significant increase in localization accuracy, with respect to the median, when compared to the traditional l^2 -norm approach. Additionally, it should be noted that the ICE approach provides the most accurate solution for all three data collects when low-quality observations are utilized.

To continue the localization performance evaluation, the localization performance of the four estimation frameworks with the high-quality observations is evaluated, as provided in Table 7.3.2. From Table 7.3.2, first, it should be noted that all four estimation frameworks are providing comparable localization statistics – as would be expected when the utilized observations adhere to the *a priori* measurement error covariance model. However, it can also be noted that the ICE approach is providing the most accurate localization statistics the majority of the time.

⁵This solution was realized with **RTKLIB**[88], which is an open-source software package for GNSS based localization.

⁶The localization performance presented within this section is significantly improved from the batch implementations presented within Chapters 5 and 6. This localization performance increase is primarily due to two modifications: 1) an accurate carrier-phase cycle slip threshold was set, 2) a static position constraint is placed on the initial and final positions within this implementation.

Table 7.3.1: Horizontal RSOS positioning error results when low fidelity receiver tracking parameters are utilized to generate the observations. The green and red cell entries correspond to the minimum and maximum statistic, respectively.

(a) Horizontal RSOS positioning error results for data collect 1.

(m.)	L_2	DCS	MM	ICE
median	2.57	0.64	1.63	0.56
std. dev.	1.41	0.98	1.05	0.72
max	10.78	9.71	10.06	13.19

(b) Horizontal RSOS positioning error results for data collect 2.

(m.)	L_2	DCS	MM	ICE
median	2.48	2.08	1.94	0.93
std. dev.	3.87	4.59	3.92	2.10
max	29.18	31.05	31.40	23.02

(c) Horizontal RSOS positioning error results for data collect 3.

(m.)	L_2	DCS	MM	ICE
median	4.41	2.82	3.62	1.48
std. dev.	2.97	3.54	3.33	5.23
max	29.53	30.38	28.30	26.61

COVARIANCE ESTIMATION ANALYSIS

To continue the evaluation, the estimated covariance from the ICE approach is assessed. Within this assessment, there are two primary objectives. First, it is desired to show that the incrementally estimated covariance represents the measurement uncertainty model. Secondly, it is desired to show that the covariance estimation process is efficiently conducted.

To enable this assessment the high-quality observations are utilized, as provided in Fig. 7.3.2. Within Fig. 7.3.2, the black points correspond to the state estimation residuals of observations which sufficiently adhere to the *a priori* measurement error uncertainty model. While, the red points correspond to the state estimation residuals of observations which were not well defined by

Table 7.3.2: Horizontal RSOS positioning error results when high fidelity receiver tracking parameters are utilized to generate the observations. The green and red cell entries correspond to the minimum and maximum statistic, respectively.

(a) Horizontal RSOS positioning error results for data collect 1.

(m.)	L_2	DCS	MM	ICE
median	0.37	0.36	0.35	0.35
std. dev.	0.30	0.27	0.29	0.28
max	5.38	5.33	5.35	5.22

(b) Horizontal RSOS positioning error results for data collect 2.

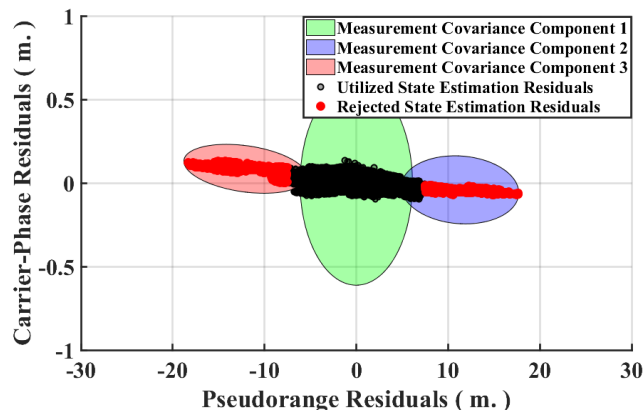
(m.)	L_2	DCS	MM	ICE
median	0.82	0.81	0.84	0.83
std. dev.	0.46	0.46	0.50	0.46
max	3.97	3.93	10.77	2.95

(c) Horizontal RSOS positioning error results for data collect 3.

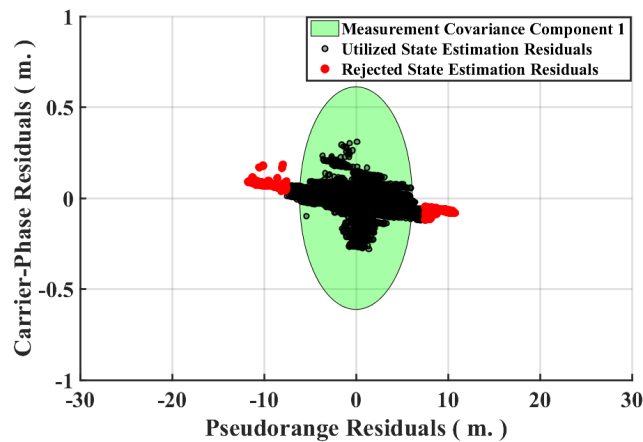
(m.)	L_2	DCS	MM	ICE
median	0.96	0.95	1.00	0.89
std. dev.	0.67	0.73	0.72	0.66
max	7.83	7.83	18.08	7.82

the *a priori* measurement uncertainty model, and thus not included during optimization; however, were utilized to modify the measurement uncertainty model. Additionally, the ellipses correspond to components of the incrementally estimated measurement error uncertainty model, with 95% confidence.

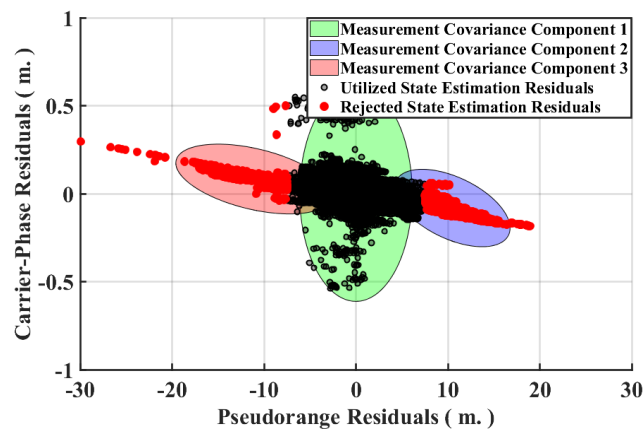
From Fig. 7.3.2, it can be seen that the incrementally estimated measurement uncertainty models closely resemble the assumed model for the high quality observations (i.e., an inlier distribution which characterizes a majority of the observations, and outlier distributions which characterize a small percentage of erroneous observations). This is specifically evident for data collects 1 and 3, as depicted in Fig. 7.3.2a and Fig. 7.3.2c, respectively.



(a) Incrementally estimated measurement error covariance model for data collect 1. For this measurement uncertainty model, approximately 91% of the observations are characterized by component 1.



(b) Incrementally estimated measurement error covariance model for data collect 2. For this data collect, only 249 observations did not adhere to the *a priori* measurement uncertainty model.



(c) Incrementally estimated measurement error covariance model for data collect 3. For this measurement uncertainty model, approximately 98% of the observations are characterized by component 1.

Figure 7.3.2: Incrementally estimated measurement error covariance model when the observations are generated with high fidelity receiver tracking parameters.

To verify the efficiency of the covariance adaptation approach, the number of times the measurement uncertainty model was adapted can be evaluated. For, data collects 1 and 3, as depicted in Fig. 7.3.2a and Fig. 7.3.2c, the covariance model was only adapted once to enable the incorporation of two outlier distributions. For data collect 2, as depicted in Fig. 7.3.2b, no covariance adaptation step was conducted – instead, only 249 observations were rejected. In contrast, if the covariance model was naively adapted every time the number of residuals were greater than the residual cardinality threshold⁷, then data collect 1 would have required 75 adaptations, data collect 2 would have required 57 adaptations, and data collect 3 would have required 91 adaptations. Thus, the incorporation of the z-test to partition the set of residuals dramatically increased the efficiency of the proposed approach.

RUN-TIME ANALYSIS

To conclude the evaluation of the proposed methodology, a run-time comparison⁸ is provided in Fig 7.3.3. From Fig. 7.3.3, it is shown that l^2 -norm, DCS, and the MM approaches all provide comparable run-time performance.

Additionally, it is clearly shown that the ICE methodology, provides the slowest average run-time; however, this slower run-time – which is still on average approximately 25 Hz – could prove to be a valid compromise when considering the significantly increase in localization accuracy granted by the approach.

Finally, although the ICE approach does currently provide the slowest run-time, an additional points should be made. For the current ICE implementation, the primary run-time bottle-neck

⁷For this study, the threshold for measurement uncertainty model adaptation, was set to 1,000 (i.e., adapt the uncertainty model if $|\mathbf{R}_O| > 1,000$).

⁸This run-time comparison was conducted on a 2.8GHz Intel Core i7-7700HQ processor.

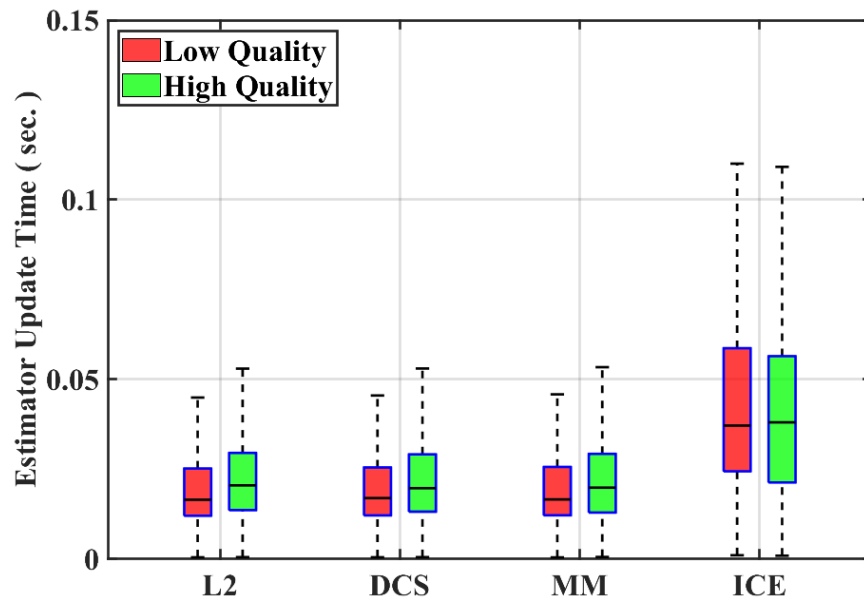


Figure 7.3.3: Estimator update time for each of the estimation frameworks over all data collects, where L_2 is a batch estimator with l^2 -norm cost function, DCS is the dynamic covariance scaling robust estimator, MM is the max-mixtures approach with a static measurement covariance model, and ICE is the proposed incremental covariance estimation technique.

for the current evaluation is implementation based. Specifically, the ICE algorithm could be implemented in such a way to dramatically decrease run-time by simply parallelizing the covariance adaptation and state estimation steps.

7.4 KEY TAKEAWAY

Within this chapter, a novel extension to the measurement uncertainty model estimation paradigm for enabling robust state estimation is proposed. Specifically, an efficient, incremental extension of the methodology is developed. The efficiency of the approach is granted by adapting the uncertainty model with only a small subset of informative state estimation residuals (i.e., the state

estimation residuals which do not adhere to the *a priori* model). The incremental nature of the approach is granted through recent advances within the probabilistic graphical model community.

To evaluate the proposed ICE approach, three degraded GNSS data sets are utilized. Based upon the results obtained on these data sets, the proposed approach provides promising results. Specifically, the proposed ICE approach provides significantly increased localization performance when utilizing degraded data, when compared to other state-of-the-art robust estimation algorithms.

8

Concluding Remarks

8.1 DISCUSSION

Several robust state estimation frameworks have been proposed over the previous decades. Underpinning all of these robust frameworks is one dubious assumption. Specifically, the assumption that an accurate *a priori* measurement uncertainty model can be provided. As systems become more autonomous, this assumption becomes less valid (i.e., as systems start operating in novel environments, there is no guarantee that the assumed *a priori* measurement uncertainty model characterizes the sensors current observation uncertainty).

In an attempt to relax this assumption, a novel robust state estimation framework is proposed. The proposed framework enables robust state estimation through the iterative adaptation of the measurement uncertainty model. The adaptation of the measurement uncertainty model is granted through non-parametric clustering of the estimators residuals, which enables the characterization of the measurement uncertainty via a Gaussian mixture model. This Gaussian mixture model based measurement uncertainty characterization can be incorporated into any non-linear least square optimization routine by only using the single assigned — the assignment of each observation to a single mode within the mixture model is provided by the utilized non-parametric clustering algorithm — component’s sufficient statistics from within the mixture model to update the uncertainty model for all observation (i.e., every observations uncertainty model is approximately characterized by the single assigned Gaussian component from within the mixture model).

Within this dissertation, the proposed framework is instantiated into three novel robust state estimation algorithms: batch covariance estimation (BCE), batch covariance estimation over an augmented data space (BCE-AD), and incremental covariance estimation (ICE). To verify the proposed framework, several GNSS data sets were collected. The collected data sets provide varying

levels of observation degradation to enable to characterization of the proposed algorithm on a diverse data set. Utilizing these data sets, it is shown that the proposed framework exhibits improved state estimation accuracy when compared to other robust estimation techniques when confronted with degraded data quality.

8.2 CONTRIBUTIONS

As discussed within Chapter 1, this dissertation is formed around three primary contributions. Now that the technical approach has been detailed and evaluated, we would like to revisit the proposed contributions.

Contribution 1: As proposed within Chapter 4, we develop a novel GNSS processing strategy. This processing techniques is built upon the connection made, within this dissertation, between the carrier-phase ambiguity states for GNSS signal processing and the loop-closure constraints which are common to the SAM formulation. By processing the GNSS signal in this way, we see two primary benefits. First, by incorporating the carrier-phase ambiguity states in this way, we can leverage the factor graph framework to exploit the sparsity of the SLAM like system (i.e., the carrier-phase ambiguity states act similar to loop-closure constraints in the typical SLAM formulation). Secondly, we see reduced time to convergence and an increase in localization accuracy, as discussed in Section 4.4.

Contribution 2: As proposed within Chapter 5 and extended within Chapter 7, a novel robust state estimation framework is developed. Specifically, we develop a novel robust state estimation framework that works through the iterative adaptation of the measurement uncertainty model. The adaptation of the measurement uncertainty model is granted through

non-parametric clustering, which enables the characterization of the measurement uncertainty via a Gaussian mixture model. The provided Gaussian mixture model can be utilized within any non-linear least squares optimization algorithm by approximately characterizing each observation with the sufficient statistics of the assigned cluster (i.e., each observation's uncertainty model is updated based upon the assignment provided by the non-parametric clustering algorithm). Within this dissertation, we develop both a batch (i.e., the batch covariance estimation (BCE), as detailed in Chapter 5) and incremental (i.e., the incremental covariance estimation (ICE), as detailed in Chapter 7) version of the proposed methodology. The proposed framework is shown to provide a significant increase in localization accuracy, when compared to other state-of-the-art algorithms, when degraded data is utilized.

Contribution 3: As proposed within Chapter 6, a novel methodology for the incorporation of metadata (e.g., the signal strength of the associated GNSS observation) into the state-estimation framework is developed. This framework (i.e., batch covariance estimation over an augmented data-space (BCE-AD)) augments the uncertainty estimation process to include metadata. The modification of the uncertainty estimation process to an augmented data space is significant because it increases the likelihood of a unique partitioning in the measurement residual domain and thus provides the ability to more accurately characterize the measurement uncertainty model. The proposed batch covariance estimation over an augmented data-space (BCE-AD) is experimentally validated on collected data where it is shown that a significant increase in state estimation accuracy can be granted compared to previously proposed robust estimation techniques.

8.3 FUTURE WORK

The work presented within this dissertation can be extended on multiple fronts. Within this section, we list a few such open questions, which are research directions that are of primary interest to the author.

8.3.1 HOW CAN WE MOST ACCURATELY ESTIMATE THE UNCERTAINTY MODEL?

The first question of interest to the author is concerned with the framework utilized to learned uncertainty model. Specifically, what framework should be utilized to find the desired pairing between computation complexity and the accuracy of the uncertainty model?

One avenue that is currently being explored is the linear response variational approach [118]. This approach is of current interest because it maintains the run-time performance of the traditional variational inference, while providing a more accurate characterization of the covariance¹. Another possible framework, which has the potential to find a desirable comprise between computation complexity and accuracy is a hybrid Markov Chain Monte Carlo variational inference approach [119].

8.3.2 HOW GENERAL ARE THE LEARNED UNCERTAINTY MODELS?

The next question of interest to the author is concerned with the generality of the learned uncertainty model. Specifically, could a model be learned for one scenario and transferred to another?

To make this more concrete, let's look at a specific example. Let's assume that we have a car driving a specific trajectory through a novel environment. While on this trajectory, a new uncertainty

¹The covariance estimate of traditional variational methods is degraded to enable increased run-time performance. This degradation is due to the mean field assumption.

model is learned through the utilization of the framework proposed within this dissertation (e.g., the ICE formulation). Now, a new car drives the same trajectory a week later. Could the previously learned uncertainty model be utilized again? If this is possible, then the computational complexity of the proposed framework could be significantly reduced through the utilization of uncertainty model sharing between platforms.

To take this one step further, we could look at the applicability of transferring uncertainty models between types of scenarios. Where, for example, a scenario type could be simply classifying all operating environments into one of three categories (e.g., rural, moderate urban, and deep urban)². If this is possible, then an *a priori* uncertainty model could be learned for each scenario type. Then, a platform could robustly localize within a novel environment with minimal covariance adaptation.

8.3.3 WHAT METADATA IS USEFUL IN OTHER DOMAINS?

As discussed within Chapter 6, the incorporation of metadata into the uncertainty model estimation process can significantly reduce the state estimation error. However, to evaluate this framework, only a single sensor modality (i.e., GNSS) was utilized. Thus, another question of interest to the author is, what metadata is useful when other sensor modalities are utilized?

To examine this question further, we would like to test several commonly used sensors (e.g., RGB, odometry, and inertial). Where, we hope to find that all of these sensor have metadata that helps to partition the measurement residual domain and thus allow for the more accurate characterization of the measurement uncertainty model, which in turn enables a more accurate state estimate.

²This partitioning of the operating environment was only selected due to its applicability to GNSS processing; however, the notation applies to any other grouping

References

- [1] Yu Gu, Jason Gross, Francis Barchesky, Haiyang Chao, and Marcello Napolitano. Avionics design for a sub-scale fault-tolerant flight control test-bed. In *Recent Advances in Aircraft Technology*. InTech, 2012.
- [2] Michael I Jordan et al. Graphical models. *Statistical Science*, 19(1):140–155, 2004.
- [3] Robin R Murphy. *Disaster robotics*. MIT press, 2014.
- [4] David Wettergreen, Greydon Foil, Michael Furlong, and David R Thompson. Science autonomy for rover subsurface exploration of the Atacama desert. *AI Magazine*, 35(4):47–60, 2014.
- [5] Laurel D Riek. Healthcare robotics. *arXiv preprint arXiv:1704.03931*, 2017.
- [6] Dan Simon. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.
- [7] E De Klerk, C Roos, and T Terlaky. *Nonlinear Optimization (CO 367)*, 2004.
- [8] Ya-xiang Yuan. A review of trust region algorithms for optimization. In *Iciam*, volume 99, pages 271–282, 2000.
- [9] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.
- [10] Dimitri P Bertsekas. Incremental least squares methods and the extended Kalman filter. *SIAM Journal on Optimization*, 6(3):807–822, 1996.
- [11] Simon J Julier and Jeffrey K Uhlmann. New extension of the Kalman filter to nonlinear systems. In *Signal processing, sensor fusion, and target recognition VI*, volume 3068, pages 182–194. International Society for Optics and Photonics, 1997.

- [12] Matthew Corwin Graham. *Robust Bayesian state estimation and mapping*. PhD thesis, Massachusetts Institute of Technology, 2015.
- [13] Frank R Hampel. Contribution to the theory of robust estimation. *Ph. D. Thesis, University of California, Berkeley*, 1968.
- [14] Peter J Huber. *Robust Statistics*. Wiley New York, 1981.
- [15] Niko Sünderhauf. *Robust optimization for simultaneous localization and mapping*. PhD thesis, Technischen Universität Chemnitz, 2012.
- [16] Pratik Agarwal, Gian Diego Tipaldi, Luciano Spinello, Cyrill Stachniss, and Wolfram Burgard. Robust map optimization using dynamic covariance scaling. In *2013 IEEE International Conference on Robotics and Automation*, pages 62–69. Citeseer, 2013.
- [17] Edwin Olson and Pratik Agarwal. Inference on networks of mixtures for robust robot mapping. *The International Journal of Robotics Research*, 32(7):826–840, 2013.
- [18] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- [19] realizing, reversing, recovering: Incremental robust loop closing over time using the irrr algorithm.
- [20] Luca Carlone, Andrea Censi, and Frank Dellaert. Selecting good measurements via l_1 relaxation: A convex approach for robust estimation over graphs. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2667–2674. IEEE, 2014.
- [21] Young C Lee et al. Analysis of range and position comparison methods as a means to provide gps integrity in the user receiver. In *Proceedings of the 42nd Annual Meeting of the Institute of Navigation*, pages 1–4, 1986.
- [22] Bradford W Parkinson and Penina Axelrad. Autonomous gps integrity monitoring using the pseudorange residual. *Navigation*, 35(2):255–274, 1988.
- [23] Mathieu Joerger, Fang-Cheng Chan, and Boris Pervan. Solution separation versus residual-based raim. *NAVIGATION: Journal of the Institute of Navigation*, 61(4):273–291, 2014.
- [24] Daniel P Shepard, Kenneth M Pesyna Jr, and Todd E Humphreys. Precise augmented reality enabled by carrier-phase differential GPS. In *Proceedings of the ION GNSS Meeting*, 2012.

- [25] Kichun Jo, Yongwoo Jo, Jae Kyu Suhr, Ho Gi Jung, and Myoungcho Sunwoo. Precise localization of an autonomous car based on probabilistic noise models of road surface marker features using multiple cameras. *IEEE Transactions on Intelligent Transportation Systems*, 16(6):3377–3392, 2015.
- [26] Paul A Rosen, Scott Hensley, Kevin Wheeler, Greg Sadowy, Tim Miller, Scott Shaffer, Ron Muellerschoen, Cathleen Jones, Howard Zebker, and Soren Madsen. UAVSAR: a new NASA airborne SAR system for science and technology research. In *Radar, 2006 IEEE Conference on*, pages 8–pp. IEEE, 2006.
- [27] Frank Dellaert, Michael Kaess, et al. Factor graphs for robot perception. *Foundations and Trends® in Robotics*, 6(1-2):1–139, 2017.
- [28] Harold W Sorenson. Least-squares estimation: from gauss to kalman. *IEEE spectrum*, 7(7):63–68, 1970.
- [29] Peter S Maybeck. *Stochastic models, estimation, and control*, volume 1. Academic press, 1982.
- [30] Gerald J Bierman. *Factorization methods for discrete sequential estimation*. Courier Corporation, 2006.
- [31] Gerald L Smith, Stanley F Schmidt, and Leonard A McGee. Application of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle. 1962.
- [32] Eric A Wan and Rudolph Van Der Merwe. The unscented kalman filter for nonlinear estimation. In *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, pages 153–158. Ieee, 2000.
- [33] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic robotics*. MIT press, 2005.
- [34] Ryan Watson, Victor Sivaneri, and Jason Gross. Performance Characterization of Tightly-Coupled GNSS Precise Point Positioning Inertial Navigation within a Simulation Environment. In *2016 AIAA Guidance Navigation and Control Conference*. AIAA, 2016.
- [35] K Ennola, J Sarvala, and G Devai. Modelling zooplankton population dynamics with the extended kalman filtering technique. *Ecological modelling*, 110(2):135–149, 1998.
- [36] Giorgio Grisetti, Rainer Kummerle, Cyrill Stachniss, and Wolfram Burgard. A tutorial on graph-based SLAM. *IEEE Intelligent Transportation Systems Magazine*, 2(4):31–43, 2010.
- [37] Peter S Maybeck. *Stochastic models, estimation, and control*, volume 2. Academic press, 1982.

- [38] Feng Lu and Evangelos Milios. Globally consistent range scan alignment for environment mapping. *Autonomous robots*, 4(4):333–349, 1997.
- [39] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.
- [40] Frank R Kschischang, Brendan J Frey, and H-A Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on information theory*, 47(2):498–519, 2001.
- [41] Frank Dellaert and Michael Kaess. Square Root SAM: Simultaneous localization and mapping via square root information smoothing. *The International Journal of Robotics Research*, 25(12):1181–1203, 2006.
- [42] Michael Kaess, Ananth Ranganathan, and Frank Dellaert. iSAM: Fast incremental smoothing and mapping with efficient data association. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 1670–1677. IEEE, 2007.
- [43] Gene H. Golub and Charles F. van Loan. *Matrix Computations*. JHU Press, fourth edition, 2013.
- [44] Michael Kaess, Viorela Ila, Richard Roberts, and Frank Dellaert. The bayes tree: An algorithmic foundation for probabilistic robot mapping. In *Algorithmic Foundations of Robotics IX*, pages 157–173. Springer, 2010.
- [45] Michael Kaess, Hordur Johannsson, Richard Roberts, Viorela Ila, John J Leonard, and Frank Dellaert. isam2: Incremental smoothing and mapping using the bayes tree. *The International Journal of Robotics Research*, 31(2):216–235, 2012.
- [46] Ryan M Watson, Jason N Gross, Clark N Taylor, and Robert C Leishman. Enabling robust state estimation through measurement error covariance adaptation. *arXiv preprint arXiv:1906.04055*, 2019.
- [47] Niko Sünderhauf and Peter Protzel. Switchable constraints for robust pose graph SLAM. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1879–1884. IEEE, 2012.
- [48] Jean RS Blair and Barry Peyton. An introduction to chordal graphs and clique trees. In *Graph theory and sparse matrix computation*, pages 1–29. Springer, 1993.
- [49] Frank Dellaert and Michael Kaess. Square Root SAM: Simultaneous localization and mapping via square root information smoothing. *The International Journal of Robotics Research*, 25(12):1181–1203, 2006.

- [50] Michael Kaess, Ananth Ranganathan, and Frank Dellaert. Fast incremental square root information smoothing. In *IJCAI*, pages 2129–2134, 2007.
- [51] Peter J Huber et al. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964.
- [52] Kirk MacTavish and Timothy D Barfoot. At all costs: A comparison of robust cost functions for camera correspondence outliers. In *2015 12th Conference on Computer and Robot Vision*, pages 62–69. IEEE, 2015.
- [53] Paul W Holland and Roy E Welsch. Robust regression using iteratively reweighted least-squares. *Communications in Statistics-theory and Methods*, 6(9):813–827, 1977.
- [54] Michael Bosse, Gabriel Agamennoni, Igor Gilitschenski, et al. Robust estimation and applications in robotics. *Foundations and Trends® in Robotics*, 4(4):225–269, 2016.
- [55] Zhengyou Zhang. Parameter estimation techniques: A tutorial with application to conic fitting. *Image and vision Computing*, 15(1):59–76, 1997.
- [56] Timothy D Barfoot. *State Estimation for Robotics*. Cambridge University Press, 2017.
- [57] Christopher Zach and Guillaume Bourmaud. Iterated lifting for robust cost optimization. In *BMVC*, 2017.
- [58] Pratik Agarwal. *Robust graph-based localization and mapping*. PhD thesis, Ph. D. dissertation, PhD thesis, University of Freiburg, Germany, 2015.
- [59] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.
- [60] JF Zumberge, MB Heflin, DC Jefferson, MM Watkins, and FH Webb. Precise point positioning for the efficient and robust analysis of GPS data from large networks. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 102(B3):5005–5017, 1997.
- [61] Pratap Misra and Per Enge. *Global Positioning System: signals, measurements and performance second edition*. Massachusetts: Ganga-Jamuna Press, 2006.
- [62] Kevin Dixon. StarFire: A global SBAS for sub-decimeter precise point positioning. In *Proceedings of ION GNSS*, pages 26–29, 2006.
- [63] Ron Muellerschoen and Yoaz Bar-Sever. Aviation applications of NASA’s global differential GPS system. In *Proceedings of the 2nd AIAA” Unmanned Unlimited” Conf. and Workshop & Exhibit, AIAA Paper*, volume 6618, 2003.

- [64] J. Gross, R. Watson, S. D’Urso, and Y. Gu. Flight-Test Evaluation of Kinematic Precise Point Positioning of Small UAVs. *International Journal of Aerospace Engineering*, 2016.
- [65] S Bisnath and Y Gao. Current state of precise point positioning and future prospects and limitations. In *Observing our changing earth*, pages 615–623. Springer, 2009.
- [66] Heroux P., Gao J., Kouba F., Lahaye Y., Mireault P., Macleod P., and Tetreault K. Products and Applications for Precise Point Positioning - Moving Towards Real-Time.
- [67] Changsheng Cai, Yang Gao, Lin Pan, and Jianjun Zhu. Precise point positioning with quad-constellations: GPS, BeiDou, GLONASS and Galileo. *Advances in space research*, 56(1):133–143, 2015.
- [68] Peter F de Bakker and Christian CJM Tiberius. Real-time multi-GNSS single-frequency precise point positioning. *GPS Solutions*, 21(4):1791–1803, 2017.
- [69] Teunissen P.J.G. and Khodabandeh A. Review and principles of PPP-RTK methods. *Journal Of Geodesy*, 89:217–240, 2015.
- [70] Bertiger W., Desai SD., Haines B., Harver N., Moore AW., Owen S., and Weiss JP. Single receiver phase ambiguity resolution with GPS data. *Journal Of Geodesy*, pages 327–337, 2010.
- [71] Groves. *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*. Artech House, 2013.
- [72] Ryan M Watson, Jason N Gross, Yoaz Bar-Sever, William I Bertiger, and Bruce J Hains. Flight data assessment of tightly coupled PPP/INS using real-time products. *IEEE Aerospace and Electronic Systems Magazine*, 32(8):10–21, 2017.
- [73] Frank R Kschischang, Brendan J Frey, and H-A Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on information theory*, 47(2):498–519, 2001.
- [74] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert. iSAM2: Incremental Smoothing and Mapping Using the Bayes Tree. *The International Journal of Robotics Research*, 31(2), 2012.
- [75] T. Whelan, M. Kaess, H. Johannsson, M. Fallon, J. Leonard, and J. McDonald. Real-Time Large-Scale Dense RGB-D SLAM With Volumetric Fusion. *The International Journal of Robotics Research*, 34(4-5), 2015.
- [76] John A Klobuchar. Ionospheric time-delay algorithm for single-frequency GPS users. *IEEE Transactions on aerospace and electronic systems*, (3):325–331, 1987.

- [77] Elliott Kaplan and Christopher Hegarty. *Understanding GPS: principles and applications*. Artech house, 2005.
- [78] B.R Schupler and T.A. Clark. How Different Antennas Affect the GPS Observable. In *GPS World*, volume 2, pages 32–36, 1991.
- [79] J-F Pascual-Sánchez. Introducing relativity in global navigation satellite systems. *Annalen der Physik*, 16(4):258–273, 2007.
- [80] P Heroux, J Kouba, P Collins, and F Lahaye. GPS carrier-phase point positioning with precise orbit products. In *Proceedings of the KIS*, pages 5–8, 2001.
- [81] Jiun-Tsong Wu, Sien C Wu, George A Hajj, Willy I Bertiger, and Stephen M Lichten. Effects of antenna orientation on GPS carrier phase. In *Astrodynamics 1991*, pages 1647–1660, 1992.
- [82] R Benjamin Harris and Richard G Mach. The GPSTk: an open source GPS toolkit. *GPS Solutions*, 11(2):145–150, 2007.
- [83] Niko Sünderhauf, Marcus Obst, Gerd Wanielik, and Peter Protzel. Multipath mitigation in GNSS-based localization using robust optimization. In *2012 IEEE Intelligent Vehicles Symposium*, pages 784–789. IEEE, 2012.
- [84] Michael Kaess, Viorela Ila, Richard Roberts, and Frank Dellaert. The Bayes Tree: An Algorithmic Foundation for Probabilistic Robot Mapping. In *WAFR*, pages 157–173. Springer, 2010.
- [85] Pinar Heggernes and Pontus Matstoms. *Finding good column orderings for sparse QR factorization*. University of Linköping, Department of Mathematics, 1996.
- [86] Pratik Agarwal and Edwin Olson. Variable reordering strategies for SLAM. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 3844–3850. IEEE, 2012.
- [87] GPSoft. *Satellite Navigation TOOLBOX 3.0 User’s Guide*. 2003.
- [88] T Takasu. RTKLIB: An open source program package for GNSS positioning, 2011.
- [89] Bela A Frigyik, Amol Kapila, and Maya R Gupta. Introduction to the Dirichlet distribution and related processes. *Department of Electrical Engineering, University of Washington, UWEETR-2010-0006*, (0006):1–27, 2010.

- [90] Persi Diaconis and Donald Ylvisaker. Conjugate priors for exponential families. *The Annals of statistics*, pages 269–281, 1979.
- [91] Stephen Tu. The Dirichlet-multinomial and Dirichlet-categorical models for Bayesian inference. *Computer Science Division, UC Berkeley*, 2014.
- [92] Kevin P Murphy. Conjugate Bayesian analysis of the Gaussian distribution. *def*, 1(2 σ^2):16, 2007.
- [93] Ignacio Alvarez, Jarad Niemi, and Matt Simpson. Bayesian inference for a covariance matrix. *arXiv preprint arXiv:1408.4050*, 2014.
- [94] Daniel Steinberg. *An unsupervised approach to modelling visual data*. PhD thesis, University of Sydney, 2013.
- [95] Matthew James Beal et al. *Variational algorithms for approximate Bayesian inference*. university of London London, 2003.
- [96] Arnaud Doucet and Xiaodong Wang. Monte Carlo methods for signal processing: a review in the statistical signal processing context. *IEEE Signal Processing Magazine*, 22(6):152–170, 2005.
- [97] David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017.
- [98] JG Liao and Arthur Berg. Sharpening Jensen’s Inequality. *The American Statistician*, pages 1–4, 2018.
- [99] Rave Harpaz and Robert Haralick. The EM algorithm as a lower bound optimization technique. *CUNY Ph. D. Program in Computer Science Technical Reports*, pages 1–14, 2006.
- [100] LabSat 3 GPS Simulator. <https://www.labsat.co.uk/index.php/en/products/labsat-3>. Accessed 3/1/19.
- [101] Carles Fernandez-Prades, Javier Arribas, Pau Closas, Carlos Aviles, and Luis Esteve. GNSS-SDR: an open source tool for researchers and developers. In *Proceedings of the 24th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2011)*, pages 780–0, 2001.
- [102] AJ Van Dierendonck, Pat Fenton, and Tom Ford. Theory and performance of narrow correlator spacing in a GPS receiver. *Navigation*, 39(3):265–283, 1992.

- [103] Werner Gurtner and Lou Estey. Rinex-the receiver independent exchange format-version 3.00. *Astronomical Institute, University of Bern and UNAVCO, Bolulder, Colorado*, 2007.
- [104] AE Niell. Global mapping functions for the atmosphere delay at radio wavelengths. *Journal of Geophysical Research: Solid Earth*, 101(B2):3227–3246, 1996.
- [105] Denis Laurichesse, Flavien Mercier, JEAN-PAUL BERTHIAS, Patrick Broca, and Luca Cerri. Integer ambiguity resolution on undifferenced GPS phase measurements and its application to PPP and satellite precise orbit determination. *Navigation*, 56(2):135–149, 2009.
- [106] Ryan M Watson and Jason N Gross. Evaluation of kinematic precise point positioning convergence with an incremental graph optimizer. In *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 589–596. IEEE, 2018.
- [107] Frank Dellaert. Factor graphs and GTSAM: A hands-on introduction. Technical report, Georgia Institute of Technology, 2012.
- [108] David M Rocke and Jian Dai. Sampling and Subsampling for Cluster Analysis in Data Mining: With Applications to Sky Survey Data. *Data Mining and Knowledge Discovery*, 7(2):215–232, 2003.
- [109] Thomas M Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE transactions on electronic computers*, (3):326–334, 1965.
- [110] Kenichi Kurihara, Max Welling, and Nikos Vlassis. Accelerated variational Dirichlet process mixtures. In *Advances in neural information processing systems*, pages 761–768, 2007.
- [111] Deng Cai, Chiyuan Zhang, and Xiaofei He. Unsupervised feature selection for multi-cluster data. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 333–342. ACM, 2010.
- [112] Tim Hesterberg, Nam Hee Choi, Lukas Meier, Chris Fraley, et al. Least angle and l_1 penalized regression: A review. *Statistics Surveys*, 2:61–93, 2008.
- [113] Jorge J Moré. The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978.
- [114] Mingzhou Song and Hongbin Wang. Highly efficient incremental estimation of Gaussian mixture models for online data stream clustering. In *Intelligent Computing: Theory and Applications III*, volume 5803, pages 174–184. International Society for Optics and Photonics, 2005.

- [115] Olivier Ledoit, Michael Wolf, et al. Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *The Annals of Statistics*, 30(4):1081–1102, 2002.
- [116] Harold Hotelling. The generalization of Student’s ratio. In *Breakthroughs in statistics*, pages 54–65. Springer, 1992.
- [117] Frank Dellaert. Factor graphs and GTSAM: A hands-on introduction. Technical report, Georgia Institute of Technology, 2012.
- [118] Ryan J Giordano, Tamara Broderick, and Michael I Jordan. Linear response methods for accurate covariance estimates from mean field variational Bayes. In *Advances in Neural Information Processing Systems*, pages 1441–1449, 2015.
- [119] Tim Salimans, Diederik Kingma, and Max Welling. Markov chain monte carlo and variational inference: Bridging the gap. In *International Conference on Machine Learning*, pages 1218–1226, 2015.