

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

INTEGRITY MONITORING WITH
NON-GAUSSIAN NOMINAL ERRORS FOR
SAFETY-CRITICAL GNSS NAVIGATION

PENGGAO YAN

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University
Department of Aeronautical and Aviation Engineering

Integrity Monitoring with Non-Gaussian Nominal Errors for
Safety-Critical GNSS Navigation

Penggao Yan

A thesis submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy
December 2024

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

Signature: _____

Name of Student: Penggao Yan

Abstract

Integrity stands as a paramount concern in civil aviation, ensuring the safety and reliability of global positioning system (GPS) navigation. With the development of new global navigation satellite system (GNSS) constellations and signals in the Aeronautical Radio Navigation Service (ARNS) bands, efforts from governments, academics, and the aviation industry are underway to promote a gradual transition from the legacy receiver autonomous integrity monitoring (RAIM) to the more advanced version, known as advanced RAIM (ARAIM). This evolution aims to facilitate a shift from primarily ensuring integrity in horizontal positioning to encompassing vertical guidance, addressing the increasingly stringent navigation requirements of modern aviation. The ARAIM algorithm has undergone regular updates over the years to incorporate new integrity analysis and performance enhancements. Nonetheless, a fundamental assumption is made in ARAIM that the nominal error is bounded by a conservative Gaussian distribution, which unnecessarily enlarges the protection level, thereby reducing the system availability under stringent navigation requirements. To release this assumption and improve the availability of integrity monitoring algorithms, this thesis prototypes a receiver autonomous integrity monitoring framework with non-Gaussian nominal errors covering GNSS error characterizing, overbounding theory, fault detection, and integrity verification.

This thesis conducts a comprehensive analysis of signal-in-space range error (SISRE) of GPS and Galileo constellations, which reveals its heavy-tailed properties. A sharp

yet conservative non-Gaussian overbound, Principal Gaussian overbound (PGO), is proposed to bound this kind of heavy-tailed error by leveraging the characteristics of the Gaussian mixture model. The overbounding property of the PGO is proved to be preserved through convolution, which makes it possible to derive measurement-level requirements from the position domain integrity requirements. Experimental results show that the PGO provides the most competitive bounding performance for SISRE when compared to the Gaussian overbound and Gaussian-Pareto overbound, yielding a sharp bound in both the core and tail parts of the error distribution. The proposed PGO served as the non-Gaussian nominal error bound for the development of fault detection and integrity monitoring algorithms in this thesis.

This thesis proposes a fault detection method, the jackknife detector, for linearized pseudorange-based positioning systems with non-Gaussian nominal error. Specifically, a test statistic based on the jackknife technique is proposed, which is proved to be the linear combination of measurement errors without any assumption about error distribution. A hypothesis test with the Bonferroni correction is constructed to detect potential faults in measurements under single-fault assumption. Then, the jackknife detector is extended to simultaneous faults by combining multiple test statistics. The reliability of the proposed method is examined in a worldwide simulation in both single- and multiple-fault settings.

This thesis proposes a multiple-hypothesis-based integrity monitoring algorithm, the jackknife ARAIM algorithm, by systematically exploiting the properties of the jackknife detector in the range domain, which is proven to be capable of handling either Gaussian or non-Gaussian nominal error bounds. A tight bound of the integrity risk is derived by quantifying the impacts of hypothetical fault vectors on the position solution. The proposed method is evaluated in a worldwide simulation with both single and dual constellations. Results reveal that the proposed method has higher system availability than the baseline ARAIM method, making it possible to support localizer performance with vertical guidance (LPV) with a decision height of 200 ft

using the GPS-Galileo dual constellation.

Publications Arising from the Thesis

1. P. Yan, Y. Hu, W. Wen, & L. T. Hsu “Multiple Faults Isolation For Multi-Constellation GNSS Positioning through Incremental Expansion of Consistent Measurements”, *IEEE Sensors Journal*, vol. 25, no. 4, pp. 6967–6981, 2025.
2. P. Yan. “Jackknife Test for Faulty GNSS Measurements Detection Under Non-Gaussian Noises”, in *Proceedings of the 37th International Technical Meeting of the Satellite Division of The Institute of Navigation*, 2024, pp. 1619–1641. (Best Student Paper Award)
3. P. Yan, Y. Zhong, & L. T. Hsu “Principal Gaussian Overbound for Heavy-tailed Error Bounding”, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 61, no. 1, pp. 829–852, 2024.
4. P. Yan, Y. Zhong, & L. T. Hsu “Bounding the Heavy-Tailed Pseudorange Error by Leveraging Membership Weights Analysis of Gaussian Mixture Model”, in *Proceedings of the ION 2024 Pacific PNT Meeting*, 2024, pp. 541–555.

Acknowledgments

First, I would like to express my sincerest gratitude to my chief supervisor, Dr. Li-Ta Hsu. Your insightful feedback and profound knowledge have illuminated the path ahead for me in academia. You have consistently encouraged me to think critically and creatively, granting me the freedom to explore the unknown. This support has significantly shaped my ability to conduct independent research. I would also like to thank my co-supervisor, Dr. Weisong Wen. Your expertise and constructive criticism have greatly enhanced the quality of my research. Thank you for your unwavering support and trust throughout my research journey.

I am also grateful to my collaborators Huang Feng, Yi Han, and Xia Xiao for their contributions. It is a highly rewarding experience to work alongside you. I am grateful for the numerous discussions we have had and the insights you have provided, which have significantly influenced the direction of this work. Furthermore, I would like to acknowledge my colleagues in the laboratory, including but not limited to Dr. Guohao Zhang, Dr. Xiwei Bai, Dr. Jiachen Zhang, Dr. Liu Jian, Jiachong Chang, Sergio Vincenzo, Xikun Liu, Max J. L. Lee, Penghui Xu, Ruizhi Hu, Liyuan Zhang, Hoi-Fung Ng, and Xi Zheng. Your enthusiasm and diverse perspectives have enhanced the impact of my research. I would also like to acknowledge my friends in the ION community. Your encouragement and constructive feedback have been invaluable throughout this journey.

Finally, I want to express my heartfelt gratitude to my family. Your support, love, and

encouragement have been my greatest source of strength throughout this challenging journey. Thank you for your patience and understanding during the long hours I dedicated to my research.

Table of Contents

Abstract	i
Publications Arising from the Thesis	iv
Acknowledgments	v
List of Figures	xii
List of Tables	xvi
Acronyms	1
Nomenclature	6
1 Introduction	16
1.1 Background and Motivation	16
1.2 Contributions	18
1.3 Thesis Outline	20
2 Fundamentals of Advanced Receiver Autonomous Integrity Moni-	

toring	22
2.1 GNSS Positioning Fundamentals	22
2.1.1 Measurement models	23
2.1.2 Error modeling of code IF combination measurements	26
2.1.3 Positioning solution	30
2.2 ARAIM Architecture	31
2.3 ARAIM Airborne Algorithm	33
2.3.1 Concepts of integrity monitoring	33
2.3.2 Range error bound	34
2.3.3 Integrity and continuity under multiple hypothesis	35
2.3.4 Solution separation based fault detection	37
2.3.5 Calculation of protection levels	39
3 Characterization of GPS/Galileo Signal-In-Space Range Error	41
3.1 Sources of Signal-In-Space Range Error	42
3.2 Determination of Orbit and Clock Errors	43
3.2.1 Data source	43
3.2.2 Antenna phase center offset correction	44
3.2.3 Computation of orbit and clock error vector	47
3.2.4 Transformation to the RAC frame	48
3.3 Definition of SISRE	48
3.4 Nominal Performance Characterization	51

3.4.1	Nominal performance of GPS SIS	52
3.4.2	Nominal performance of Galileo SIS	55
4	Principal Gaussian Overbound for Heavy-tailed Error Distribution	60
4.1	Review of Overbounding Methods	61
4.1.1	Two-step Gaussian overbound	62
4.1.2	Gaussian-Pareto overbound	63
4.2	Gaussian Mixture Model	65
4.3	Principal Gaussian Overbound for the Heavy-Tailed Error Distribution	66
4.3.1	Membership weight analysis	66
4.3.2	Dominance partition	68
4.3.3	Principal Gaussian overbound	72
4.3.4	Preservation of overbounding property	79
4.3.5	Numerical consideration of convolution	80
4.4	Bounding Performance of SISRE	84
4.4.1	Two-side heavy-tailed SISRE	85
4.4.2	One-side heavy-tailed SISRE	87
4.4.3	Gaussian-like SISRE	88
4.4.4	Bounding parameters for individual satellites	89
4.5	Non-Gaussian Nominal Error Bounds of Code IF Combination	90
5	Fault Detection with Non-Gaussian Nominal Error Bound	94
5.1	Jackknife Detector for Single Fault Detection	95

5.1.1	Full set and subset solutions based on weighted least square	95
5.1.2	Construction of jackknife residual	96
5.1.3	Distribution of jackknife residual	97
5.1.4	Jackknife test for fault detection	99
5.2	Extend Jackknife Detector to Simultaneous Faults	100
5.2.1	Reconstruction of jackknife residual	100
5.2.2	Combination of jackknife residuals	101
5.2.3	Reconstruction of hypothesis tests	102
5.3	Detection Performance with Worldwide Simulations	103
5.3.1	Nominal error simulation and bounding	105
5.3.2	Single-fault detection performance	107
5.3.3	Multiple-fault detection performance	108
6	Integrity Monitoring with Non-Gaussian Nominal Error Bound	111
6.1	Threat Model and Multiple Hypotheses	112
6.2	Determine the Threshold of Monitors	112
6.3	Integrity Risk Evaluation	113
6.3.1	Bound on the probability of HMI under H_0	114
6.3.2	Bound on the probability of HMI under H_k	115
6.3.3	Finalized bound of integrity risk	118
6.4	Protection Level Derivation	119
6.5	Consideration of Constellation Faults	120

6.6	Worldwide Simulation	122
6.6.1	Single-constellation experiments	124
6.6.2	Dual-constellation experiments	128
7	Conclusion	132
7.1	Summary of Accomplishments	132
7.2	Recommended Topics for Future Research	135
7.3	Closing	137
A	Appendices	138
A.1	Gaussian Overbound of Multipath and Code Noise	138
A.2	Maximum Number of Simultaneous Faults That Need to be Monitored	139
A.3	The EM Algorithm	141
A.4	Proof of Tail Region Overbound	142
A.5	Proof of Monotonicity	144
A.6	Distribution of Linear Combination of Random Variables	145
A.7	Distribution of Jackknife Residual under Gaussian Noises	146
A.8	Bonferroni Correction	148
	Bibliography	150

List of Figures

2.1	Illustration of the multiple hypothesis solution separation in the base-line ARAIM airborne algorithm.	36
3.1	Different conventions for orbit and clock reference points in GNSS broadcast and precise ephemerides. The green dot represents the APC used in BCE and the blue dot represents the APC used in PCE. . . .	45
3.2	Definition of the BF frame of satellites.	46
3.3	Demonstration of orbit error.	47
3.4	Definition of satellite RAC frame.	49
3.5	Demonstration of IURE over a mesh of 200 user locations.	49
3.6	The folded CDF of GPS radial, along-track, cross-track, clock, and $SISRE_{UPE}$ errors from January 1st, 2020 to December 31st, 2022 . .	53
3.7	The folded CDF of GPS $SISRE_{UPE}$ for individual satellites from January 1st, 2020 to December 31st, 2022.	53
3.8	The QQ plot of GPS $SISRE_{UPE}$ for individual satellites from January 1st, 2020 to December 31st, 2022. The x-axis represents the quantile of error distribution (unit: meter) while the y-axis stands for the standard normal quantile (unit: meter).	54

3.9	The folded CDF of Galileo radial, along-track, cross-track, clock, and $SISRE_{UPE}$ errors from January 1st, 2020 to December 31st, 2022. . .	56
3.10	The folded CDF of Galileo $SISRE_{UPE}$ for individual satellites from January 1st, 2020 to December 31st, 2022.	56
3.11	The QQ plot of Galileo $SISRE_{UPE}$ for individual satellites from January 1st, 2020 to December 31st, 2022. The x-axis represents the quantile of error distribution (unit: meter) while the y-axis stands for the standard normal quantile (unit: meter).	57
4.1	(a) Membership weights and (b) the relative kurtosis error of a zero-mean BGMM with $p_1 = 0.9$, $\sigma_1 = 0.5$, and $\sigma_2 = 0.7$; (c) Membership weights and (d) the relative kurtosis error of a zero-mean BGMM with $p_1 = 0.9$, $\sigma_1 = 0.5$, and $\sigma_2 = 1.5$. x_{lp} and x_{rp} are the core-tail transition points. The two intersection points of $s_1(x)$ and $s_2(x)$ are marked with blue and red solid points.	69
4.2	The (a) PDF and (b) CDF of the Principal Gaussian overbound of a zero-mean bimodal Gaussian mixture model with $p_1 = 0.9$, $\sigma_1 = 0.5$, and $\sigma_2 = 1.5$. The two-step Gaussian overbound is plotted for comparison.	76
4.3	Demonstrations of (a) the discretization process and (b) the quantile searching process. The shaded area in (b) shows the cumulative probability from t_1 to t_{m-1}	83
4.4	The folded CDF of $SISRE_{UPE}$ and its bounding results for (a) GPS satellite SVN63; and (b) Galileo satellite GSAT0206.	86
4.5	The folded CDF of $SISRE_{UPE}$ and its bounding results for (a) GPS satellite SVN66; and (b) Galileo satellite GSAT0212.	87

4.6	The folded CDF of $SISRE_{UPE}$ and its bounding results for GPS satellite SVN46.	89
5.1	The setting of the nominal error and bound of the code IF combination. ‘ob’ represents ‘Gaussian’, ‘acc’, or ‘int’.	105
5.2	Detection rate contour of (a) the SS detector, (b) JKD-Gaussian, and (c) JKD-non-Gaussian with single artificially injected bias. (d) The histogram of the detection rate difference between the JKD-non-Gaussian and the SS detector.	107
5.3	Detection rate contour of (a) the SS detector, (b) JKD-Gaussian, and (c) JKD-non-Gaussian with two artificially injected biases. (d) The histogram of the detection rate difference between the JKD-non-Gaussian and the SS detector.	109
6.1	99.5 percentile of the VPL over the course of the day yielded by (a) the baseline ARAIM and (b) the proposed JK-Gaussian ARAIM for the single constellation; and the triangular chart of (c) the baseline ARAIM and (d) the proposed JK-Gaussian ARAIM regarding the vertical performance for the single constellation. “NO” represents normal operation, “MI” represents misleading information, “SU” represents system unavailable, “SU&MI” represents system unavailable and misleading information, and “HMI” represents hazardously misleading information.126	
6.2	(a) 99.5 percentile of the VPL over the course of the day yielded by the proposed JK-non-Gaussian ARAIM for the single constellation; (b) The triangular chart of the proposed JK-non-Gaussian ARAIM regarding the vertical performance for the single constellation.	127

6.3	99.5 percentile of the VPL over the course of the day yielded by (a) the baseline ARAIM and (b) the proposed JK-Gaussian ARAIM for the dual constellation; and the triangular chart of (c) the baseline ARAIM and (d) the proposed JK-Gaussian ARAIM regarding the vertical performance for the dual constellation.	129
6.4	(a) 99.5 percentile of the VPL over the course of the day yielded by the proposed JK-non-Gaussian ARAIM for the dual constellation; (b) The triangular chart of the proposed JK-non-Gaussian ARAIM regarding the vertical performance for the dual constellation.	131

List of Tables

3.1	GPS fault events between 2020 and 2022 (Taken from [1])	51
3.2	Galileo fault events between 2020 and 2022 (Taken from [1])	51
3.3	The mean and standard deviation of the radial, along-track, cross-track, clock, and $SISRE_{UPE}$ errors for individual GPS satellites from 1/1/2020 to 12/31/2022 (unit: cm)	58
3.4	The mean and standard deviation of the radial, along-track, cross-track, clock, and $SISRE_{UPE}$ errors for individual Galileo satellites from 1/1/2020 to 12/31/2022 (unit: cm)	59
4.1	Parameters of the Gaussian overbound and the Principal Gaussian overbound of $SISRE_{UPE}$ for each GPS satellite (unit: m)	92
4.2	Parameters of the Gaussian overbound and the Principal Gaussian overbound of $SISRE_{UPE}$ for each Galileo satellite (unit: m)	93
5.1	Source of almanacs of the GPS and Galileo constellations	104
5.2	Parameters of the fault detection experiments	104
5.3	Overbounds used in different detectors	106
6.1	Overbounds used in different ARAIM algorithms	124

6.2	Parameters used for evaluating integrity monitoring algorithms in the simulation	125
6.3	Coverage for the single constellation at different levels of system availability	128
6.4	Coverage for the dual constellation at different levels of system availability	131
A.1	The code noise and multipath error bound for Galileo airborne receiver against the elevation angle	139

Acronyms

AL	Alert Limit.
ANTEX	ANTenna EXchage.
APC	Antenna Phase Center.
ARAIM	Advanced Receiver Autonomous Integrity Monitoring.
ARNS	Aeronautical Radio Navigation Service.
BCE	Broadcast Ephemeris.
BF	Body Fixed.
BGMM	Bimodal Gaussian Mixture Model.
C/A	Coarse/Acquisition.
CDF	Cumulative Distribution Function.
CODE	Center for Orbit Determination in Europe.
CoM	Center of Mass.
CSP	Constellation Service Providers.
DCB	Differential Code Bias.
DFT	Discrete Fourier Transform.
DOP	Dilution Of Precision.

Acronyms

ECEF	Earth-centered Earth-fixed.
EM	Expectation-Maximization.
F/NAV	Freely Accessible Navigation Message.
FDE	Fault Detection and Exclusion.
FT	Fourier Transform.
GBAS	Ground-based Augmentation Systems.
GLONASS	GLObalnaya NAvigatsionnaya Sputnikovaya Sistema in Russian.
GMM	Gaussian Mixture Model.
GNSS	Global Navigation Satellite System.
GPS	Global Positioning System.
GPS-SPS-PS	GPS Standard Positioning Service Performance Standard.
GSA	European GNSS Agency.
HMI	Hazardously Misleading Information.
HPL	Horizontal Protection Level.
I/NAV	Integrity Navigation Message.
IDFT	Inverse Discrete Fourier Transform.
IF	Ionosphere Free.
IFT	Inverse Fourier Transform.
IGS	International GNSS Service.
ISM	Integrity Support Message.

IURE	Instantaneous User Range Error.
JKD-non-Gaussian	Jackknife Detector using Non-Gaussian Nominal Overbound.
JKD-Gaussian	Jackknife Detector using Gaussian Nominal Overbound.
LOS	Line-of-Sight.
LPV	Localizer Performance with Vertical Guidance.
MAAST	MATLAB Algorithm Availability Simulation Tool.
MCS	Master Control Station.
MGEX	Multi-GNSS EXperiment.
MHSS	Multiple Hypothesis Solution Separation.
MI	Misleading Information.
MLE	Maximum Likelihood Estimation.
MOPS	Minimum Operational Performance Standard.
NGA	National Geospatial-Intelligence Agency.
ODTS	Orbit Determination and Time Synchronization.
OS-SDD	Open Service Service Definition Document.
PCE	Precise Ephemeris.
PDF	Probability Density Function.

Acronyms

PGO	Principal Gaussian Overbound.
PL	Protection Level.
PMF	Probability Mass Function.
QQ	Quantile-Quantile.
QZSS	Quasi-Zenith Satellite System.
RAC	Radial, Along-Track and Cross-Track.
RAIM	Receiver Autonomous Integrity Monitoring.
RINEX	Receiver Independent Exchange.
RTCA	Radio Technical Commission for Aeronautics.
SBAS	Satellite-based Augmentation Systems.
SIS	Signal-in-Space.
SISRE	Signal-in-Space Range Error.
SS	Solution Separation.
TTA	Time to Alert.
UPE	User Projected Error.
URA	User Range Accuracy.
VAL	Vertical Alert Limit.
VPE	Vertical Positioning Error.
VPL	Vertical Protection Level.
WAAS	Wide Area Augmentation System.

WGC Worldwide GNSS Committee.

WLS Weighted Least Square.

Nomenclature

Measurements

$\varrho_{j,f}^i$	Pseudorange measurement regarding the satellite i and the receiver j at frequency f
$\rho_{j,f}^i$	Pseudorange measurement with tropospheric and clock offset corrections regarding the satellite i and the receiver j at frequency f
$\varrho_{j,IF}^i$	Code IF combination measurement regarding the satellite i and the receiver j
r_j^i	Geometry distance between the satellite i and the receiver j
$\varphi_{j,f}^i$	Carrier phase measurement regarding the satellite i and the receiver j at frequency f
τ_j	Clock offset of the receiver j from the GNSS time scale
τ^i	Clock offset of the satellite i from the GNSS time scale
$c\delta_j$	Clock offset of the receiver j regarding the code IF combination
$c\delta^i$	Clock offset of the satellite i regarding the code IF combination
Td_j^i	Tropospheric delay regarding the satellite i and the receiver j

$Id_{j,f}^i$	Ionospheric delay regarding the satellite i and the receiver j at frequency f
$K_{\varrho,j,f}$	Code instrumental delay of the receiver j at frequency f
$K_{\varrho,f}^i$	Code instrumental delay of the satellite i at frequency f
$K_{\varphi,j,f}$	Carrier phase instrumental delay of the receiver j at frequency f
$K_{\varphi,f}^i$	Carrier phase instrumental delay of the satellite i at frequency f
$K_{\varrho,j,IF}$	DCB of the receiver j regarding frequencies f_A and f_B
$K_{\varrho,IF}^i$	DCB of the satellite i regarding frequencies f_A and f_B
$n_{j,f}^i$	Ambiguity term of the carrier phase measurement regarding the satellite i and the receiver j at frequency f

Measurement Errors

$\varepsilon_{\varrho,j,f}^i$	Residual errors of the pseudorange measurement regarding the satellite i and the receiver j at frequency f
$\varepsilon_{\varphi,j,f}^i$	Residual errors of the carrier phase measurement regarding the satellite i and the receiver j at frequency f
$\varepsilon_{orb\&clk}^i$	Ephemeris and clock error of satellite i
$\varepsilon_{tropo,j}^i$	Residual tropospheric error regarding the satellite i and the receiver j
$\varepsilon_{iono,j,f}^i$	Residual ionospheric error regarding the satellite i and the receiver j at frequency f
$\varepsilon_{\varrho,user,j,f}^i$	Pseudorange multipath and code noises regarding the satellite i and the receiver j at frequency f

Nomenclature

$\varepsilon_{\varphi,user,j,f}^i$	Carrier phase multipath and code noises regarding the satellite i and the receiver j at frequency f
$\varepsilon_{\varrho,j,IF}^i$	Residual errors of the code IF combination measurement regarding the satellite i and the receiver j
$\varepsilon_{\varrho,user,j,IF}^i$	Multipath and code noises of the code IF combination regarding the satellite i and the receiver j

Measurement Error Bounds

$\sigma_{orb\&clk}^i$	Standard deviation of the orbit and clock error bound regarding the satellite i
$\sigma_{tropo,j}^i$	Standard deviation of the residual tropospheric error bound regarding the satellite i and the receiver j
$\sigma_{\varrho,user,j,IF}^i$	Standard deviation of the multipath and code noise error bound regarding the satellite i and the receiver j in the IF combination of the code measurement model
$\sigma_{\varrho,j,IF}^i$	Standard deviation of the Gaussian bound of $\varepsilon_{\varrho,j,IF}^i$
θ_j^i	Elevation angle associated with the receiver j and the satellite i
$f_{orb\&clk}^i(\cdot)$	PDF of the range projection of clock and orbit error for satellite i
$f_{tropo,j}^i(\cdot)$	PDF of the tropospheric error regarding satellite i and receiver j
$f_{\varrho,user,j,IF}^i(\cdot)$	PDF of the multipath and code noise of code IF combination measurement regarding satellite i and receiver j
$f_{orb\&clk,PGO}^i(x)$	PDF of the PGO of the range projection of clock and orbit error for satellite i
$f_{tropo,j,ob}^i(x)$	PDF of the Gaussian overbound of the tropospheric error

$f_{\varrho, user, j, IF, ob}^i(x)$	PDF of the Gaussian overbound of the multipath and code noise
$f_{\varrho, j, IF, Gaussian}^i(\cdot)$	PDF of the Gaussian overbound for code IF combination measurement regarding satellite i and receiver j
$f_{\varrho, j, IF, acc}^i(\cdot)$	PDF of the non-Gaussian overbound for code IF combination measurement regarding satellite i and receiver j for accuracy evaluation and fault detection purposes
$G_{\varrho, j, IF, acc}^i(x)$	CDF of the non-Gaussian overbound for code IF combination measurement regarding satellite i and receiver j for accuracy evaluation and fault detection purposes
$G_{\varrho, j, IF, int}^i(\cdot)$	CDF of the non-Gaussian overbound for code IF combination measurement regarding satellite i and receiver j for integrity purposes

Position, velocity, and timing

$\mathbf{p}^i$	$\mathbf{p}^i = [p_x^i, p_y^i, p_z^i]$ is the position of the i th satellite in the ECEF frame
$\mathbf{u}_j$	$\mathbf{u}_j = [u_{j,x}, u_{j,y}, u_{j,z}]$ is the position of the j th receiver in the ECEF frame
$\mathbf{x}_t$	Receiver true state vector
$\mathbf{x}_0$	Linearization point
$\mathbf{x}$	System state vector
N_{const}	Number of constellations

Estimation

n	Number of measurements
m	Size of the receiver state. $m = 3 + N_{const}$

Nomenclature

$\mathbf{y}$	Measurement vector
$\boldsymbol{\varepsilon}$	Measurement error vector
$\hat{\mathbf{x}}$	Full set solution
$\hat{\mathbf{x}}^{(k)}$	Solution for subset k , where $k = 0$ corresponds to the full set
$\hat{\mathbf{x}}_t^{(k)}$	Estimation of the positioning state $x_t^{(k)}$ associated with the k th subset
$\mathbf{G}$	Geometry matrix for full set measurements
$\mathbf{S}$	Solution matrix for full set measurements
$\mathbf{W}$	Weight matrix for full set measurements
$\mathbf{S}^{(k)}$	Solution matrix for subset k , where $k = 0$ corresponds to the full set
$\mathbf{W}^{(k)}$	Weight matrix for subset k , where $k = 0$ corresponds to the full set
$\mathbf{G}^{(k,*)}$	Geometry matrix associated with the k th jackknife test statistic. The dimension is $(n - 1) \times m$
$\mathbf{S}^{(k,*)}$	Solution matrix associated with the k th jackknife test statistic. The dimension is $m \times (n - 1)$
$\tilde{\mathbf{P}}^{(k)}$	Reconstructed Projection matrix related to the k th jackknife test statistic
ε_i	The i th element of $\boldsymbol{\varepsilon}$
$\mathbf{g}_k$	The k th row of $\mathbf{G}$
$S_{v,i}$	The (v, i) th element in $\mathbf{S}$
$\tilde{\mathbf{p}}_k$	The k th row of $\mathbf{I} - \tilde{\mathbf{P}}$

$\tilde{p}_{k,j}$ The j th element of $\tilde{\mathbf{p}}_k$

Integrity and continuity budgets

I_{REQ} Integrity requirement of the system

I_{REQ}^v I_{REQ} allocated to v th component. I_{REQ}^3 is the allocated budget for the vertical component, and $I_{\text{REQ}}^1 + I_{\text{REQ}}^2$ is the allocated budget for the horizontal component

$C_{\text{REQ,FA}}$ Continuity budget allocated to false alarms

$C_{\text{REQ,FA}}^v$ $C_{\text{REQ,FA}}$ allocated to v th component. $C_{\text{REQ,FA}}^3$ is the allocated budget for the vertical component, and $C_{\text{REQ,FA}}^1 + C_{\text{REQ,FA}}^2$ is the allocated budget for the horizontal component

General notations for estimation and inference

P_{HMI} Integrity risk

$P_{\text{continuity}}$ Continuity risk

P_{sat} Prior probability of satellite fault per approach

P_{const} Prior probability of constellation fault per approach

P_{THRES} Threshold for the integrity risk coming from unmonitored faults

P_{H_k} Prior probability of fault mode k

$P_{\text{not monitored}}$ Prior probability of the unmonitored events

$P_{\text{Fault-free}}$ Prior probability of no fault

k_{max} Maximum number of simultaneous faults that need to be monitored

$N_{\text{fault modes}}$ The finally solved $N_{\text{fault,max}}$

Nomenclature

e_0	Estimation error on the parameter of interest
ℓ	The specified alert limit
ℓ_v	Alert limit of the v th positioning component
q_k	General test statistic that associated with the test H_0 v.s. H_k
R_k	Rejection region of q_k
idx_k^{ex}	Set of indices of faulty measurements in fault model k
PL_v	Protection level of the v th positioning component. PL_3 is the VPL, and $\sqrt{PL_1^2 + PL_2^2}$ is the HPL
$\mathbf{e}_v$	A column vector with v th entry to be 1 and all others to be 0

Estimation and inference in baseline ARAIM

σ_{URE}^i	1-sigma orbit and clock error bound of the satellite i for accuracy
σ_{URA}^i	1-sigma orbit and clock error bound of the satellite i for integrity
b_{nom}^i	Nominal bias for integrity regarding the satellite i
$\sigma_{acc,j}^i$	1-sigma accuracy error bound regarding the satellite i and the receiver j
$\sigma_{int,j}^i$	1-sigma integrity error bound regarding the satellite i and the receiver j
$\mathbf{C}_{acc}$	Covariance matrix for accuracy
$\mathbf{C}_{int}$	Covariance matrix for integrity
$\sigma_{ss,v}^{(k)}$	1-sigma bound of $d_v^{(k)}$
$\sigma_v^{(k)}$	1-sigma bound of the positioning solution $\hat{\mathbf{x}}^{(k)}$

$b_v^{(k)}$	The worst-case impact of the nominal biases of the integrity range error bound on the position solution $\hat{\mathbf{x}}^{(k)}$
$d_v^{(k)}$	Test statistic regarding fault mode k in the solution separation method
$D_{k,v}$	Threshold of the solution separation integrity monitor
$Q^{-1}(\cdot)$	Quantile function of the standard normal distribution

Estimation and inference in Jackknife detector and ARAIM

$\mathbf{y}^{(k,*)}$	The $\mathbf{y}$ vector associated with fault mode k . The dimension is $(n - 1) \times 1$
$\hat{\mathbf{y}}^{(k,*)}$	The predicted measurement vector based on subsolution $\hat{\mathbf{x}}$
$\tilde{\mathbf{y}}^{(k)}$	The reconstructed $\mathbf{y}$ vector associated for fault mode k . The dimension is $n \times 1$
$\hat{y}_k$	Predicted k th measurement based on subsolution $\hat{\mathbf{x}}^{(k)}$
y_k	The k th element of $\mathbf{y}$
$\boldsymbol{\varepsilon}^{(k,*)}$	The $\boldsymbol{\varepsilon}$ vector associated with fault mode k . The dimension is $(n - 1) \times 1$
ε_k	The k th element of $\boldsymbol{\varepsilon}_k$
t_k	The Jackknife residual (detector) for the k th subset under single-fault hypothesis
$\tilde{t}_k$	The Jackknife residual (detector) for the k th subset under multi-fault hypothesis
t_k^*	The unified Jackknife residual (detector) for the k th subset
T_k	Threshold for the Jackknife detector t_k^*

τ	Upper limit of the significance level for the multiple testing problem
$\mathbf{b}^{(k)}$	The hypothetical fault vector in hypothesis H_k
$b_j^{(k)}$	The j th element of $\mathbf{b}^{(k)}$
$\mathbf{q}^{(k)}$	Coefficient vector for constructing the positioning error in hypothesis H_k
$f_{(\hat{\mathbf{x}}-\mathbf{x})_v}(\cdot)$	The PDF of $(\hat{\mathbf{x}} - \mathbf{x})_v$
$f_{\mathbf{q}^{(k)}\boldsymbol{\varepsilon}}(\cdot)$	The PDF of $\mathbf{q}^{(k)}\boldsymbol{\varepsilon}$
$Q_{t_k^*}^{-1}(\cdot)$	Quantile function of the distribution of t_k^*
$Q_{(\hat{\mathbf{x}}-\mathbf{x})_v}^{-1}(\cdot)$	The quantile functions of $(\hat{\mathbf{x}} - \mathbf{x})_v$
$Q_{\mathbf{q}^{(k)}\boldsymbol{\varepsilon}}^{-1}(\cdot)$	The quantile functions of $\mathbf{q}^{(k)}\boldsymbol{\varepsilon}$

Satellite Ephemeris and Clock Errors

$\mathbf{x}_{CoM,ECEF}^{i,k,BCE}$	The broadcast CoM position of satellite i at epoch k in the ECEF coordinate system
$\mathbf{x}_{APC,ECEF}^{i,k,BCE}$	The broadcast APC position of satellite i at epoch k in the ECEF coordinate system
$\mathbf{a}_{BCE}^i$	The APC offset vector of satellite i defined in the satellite BF frame for the BCE. $\mathbf{a}_{BCE}^i = [a_{x,BCE}^i, a_{y,BCE}^i, a_{z,BCE}^i]^T$
$\mathbf{a}_{PCE}^i$	The APC offset vector of satellite i defined in the satellite BF frame for the PCE. $\mathbf{a}_{PCE}^i = [a_{x,PCE}^i, a_{y,PCE}^i, a_{z,PCE}^i]^T$
$\mathbf{R}_{BF}^{i,k,ECEF}$	The rotation matrix from the BF frame to the ECEF frame for satellite i at epoch k
$\mathbf{x}_{sat\ to\ sun}^{i,k}$	Satellite-sun vector regarding satellite i at epoch k

$\mathbf{x}_{sun}^k$	Sun position at epoch k in the ECEF coordinate system
$\mathbf{p}^{i,k}$	Position of satellite i at epoch k in the ECEF coordinate system
$c\delta_{CoM,BCE}^{i,k}$	CoM-referenced clock offset of satellite i at epoch k in the BCE
$c\delta_{APC,BCE}^{i,k}$	APC-referenced clock offset of satellite i at epoch k in the BCE
$c\delta_{CoM,PCE}^{i,k}$	CoM-referenced clock offset of satellite i at epoch k in the PCE
$c\delta_{APC,PCE}^{i,k}$	APC-referenced clock offset of satellite i at epoch k in the PCE
$\boldsymbol{\xi}_{orb,ECEF}^{i,k}$	Orbit error vector of satellite i at time epoch k in the ECEF coordinate system
$\xi_{clk,ECEF}^{i,k}$	The CoM-referenced clock error of satellite i at time epoch k
$\tilde{\xi}_{clk,ECEF}^{i,k}$	The CoM-referenced clock error of satellite i at time epoch k with constellation mean correction
$\boldsymbol{\xi}_{orb,RAC}^{i,k}$	Orbit error vector of satellite i at time epoch k in the RAM frame. $\boldsymbol{\xi}_{orb,RAC}^{i,k} = [\xi_{orb,R}^{i,k}, \xi_{orb,A}^{i,k}, \xi_{orb,C}^{i,k}]$
$\mathbf{R}_{ECEF}^{i,k,RAC}$	The rotation matrix from the ECEF frame to the RAC frame for satellite i at epoch k
$\mathbf{v}^{i,k}$	Velocity of satellite i at epoch k in the ECEF frame
$IURE_{j,orb}^{i,k}$	Orbit component of the IURE regarding satellite i and receiver j at epoch k
$IURE_j^{i,k}$	IURE regarding satellite i and receiver j at epoch k
$UPE^{i,k}$	UPE of satellite i at epoch k
$SISRE_{UPE}^i$	SISRE of satellite i over a period

Chapter 1

Introduction

1.1 Background and Motivation

Safety-of-life systems, such as satellite-based augmentation systems (SBAS), ground-based augmentation systems (GBAS), and receiver autonomous integrity monitoring (RAIM), are important in modern aviation to ensure the safety and reliability of navigation systems [2, 3]. These systems are designed to provide integrity-assured position solutions to GNSS users and typically require a low integrity risk. Among these safety-of-life systems, RAIM has been a significant focus in the aviation industry over the last thirty years.

The legacy RAIM was first introduced in the 1980s and was designed to detect measurement faults and provide error bounds for lateral navigation with the global positioning system (GPS) [4–10]. In the early exploration, the range-comparison method [4], the least-squared-residuals method [8], and the parity method [9] have been proposed, which establish the basis of autonomous integrity monitoring. The equivalence relationship among the three methods was found in [3], which established the baseline RAIM scheme. Later, Walter and Enge [2] proposed the weighted form for RAIM algorithms, which improves the accuracy and integrity of the navigation

solutions. However, these RAIM algorithms face limitations in precise vertical navigation due to factors like uncertainties in the ionospheric model and high dilution of precision (DOP).

The modernization of GPS [11–13] and the emerging GNSS constellations present the chance to elevate legacy RAIM to advanced RAIM (ARAIM), which aims to support more robust and precise aircraft guidance for en route, terminal, and approach operations in civil aviation [14]. Pervan et al. [15] developed the multiple hypothesis solution separation (MHSS) method to directly evaluate integrity risks under the unified consideration of all single-element failure hypotheses and the no-failure hypothesis, which lays the foundation for the current airborne ARAIM algorithm. Later, Blanch et al. [16] expanded this work by optimally allocating the integrity and continuity budget among the failure modes to obtain the minimum protection level. In parallel, Joerger et al. developed the residual-based ARAIM [17] and revealed its equivalent relationship with the solution separation-based ARAIM [18]. The detailed comparison of the residual-based ARAIM and solution-separation ARAIM in [18] provides a practical guideline for choosing proper algorithms in different applications. These works together contributed to the standardization of the baseline ARAIM algorithm [19] and its subsequent refinement based on the latest safety analyses [20, 21].

The baseline ARAIM algorithm [20] provides a standard procedure to ensure the safety and reliability of navigation systems, mainly consisting of range error modeling, fault detection, and protection level (PL) calculation. Specifically, the Gaussian model is employed to bound nominal range errors, including the range projection of clock and orbit errors, tropospheric error, and multipath and code noise errors. Then, a set of test statistics, named solution separation, is constructed by comparing the navigation solutions under the fault-free hypothesis with many other alternative hypotheses. By leveraging the Gaussian properties of the nominal error bounds, the threshold for each solution separation test statistic is derived to detect potential faulty measurements in the system. Finally, the bound of the positioning solution error that

satisfies integrity requirements, namely, the PL, is calculated. The whole algorithm is built on the Gaussian overbounds for nominal range errors, which simplifies the derivation and reduces the computational effort. However, nominal range errors in the real world usually have non-Gaussian and heavy-tailed properties [22–24]. For example, as important components of range errors, orbit and clock errors show significantly heavy-tailed properties [25, 26], making their Gaussian overbound over-conservative. Such over-conservatism will be passed to the position domain and enlarge the PL of the baseline ARAIM algorithm, eventually hindering the system’s availability in real-world applications under stringent navigation requirements, such as the LPV-200 precision approach [27] and urban air mobility [28].

1.2 Contributions

To address the conservatism issue in the baseline ARAIM algorithm, this thesis systematically incorporates non-Gaussian modeling into the construction of the integrity monitoring algorithm, which results in the following five contributions.

(1) Discover heavy-tailed properties of GPS/Galileo signal-in-space range errors

This thesis characterizes the nominal performance of signal-in-space range error (SISRE) for GPS and Galileo satellites using broadcast and precise ephemerides data from 2020–2022, which reveals the heavy-tailed properties of SISRE. Since the SISRE is one of the most important parts of range errors, the results provide empirical justification for developing non-Gaussian overbounding methods and incorporating non-Gaussian modeling into integrity monitoring methods. Results also show that the Galileo SISRE has a smaller standard deviation but exhibits heavier tails than the GPS SISRE, which plays an important role in explaining the performance change of integrity monitoring algorithms in different constellation settings.

(2) Construct a sharp yet conservative overbound for heavy-tailed error distributions

This thesis proposes the Principal Gaussian overbound (PGO) for heavy-tailed error distributions by leveraging the characteristics of the Gaussian mixture model (GMM). The overbounding property of the PGO is proved to be preserved through convolution, which makes it possible to derive pseudorange-level requirements from the position domain integrity requirements. The proposed PGO is then applied to tightly bound the GPS/Galileo SISRE, based on which, the non-Gaussian nominal error bounds of code ionosphere-free (IF) combination are constructed, providing nominal error models for constructing non-Gaussian fault detection and integrity monitoring algorithms.

(3) Develop a fault detection method with non-Gaussian nominal errors

This thesis develops the jackknife detector, which provides the theoretical foundation for detecting single faults in linearized pseudorange-based positioning systems under non-Gaussian nominal errors. The proposed fault detection method is further extended to detect simultaneous faults by combining multiple test statistics. In a worldwide simulation, the proposed fault detection method demonstrates superior performance than the solution separation fault detection method under non-Gaussian nominal errors. The proposed fault detection method provides foundations to improve the availability of integrity monitoring algorithms under stringent navigation requirements.

(4) Prototype an integrity monitoring algorithm with high availability under stringent navigation requirements

This thesis proposes the jackknife ARAIM, a multiple-hypothesis-based integrity monitoring algorithm, capable of handling either Gaussian or non-Gaussian nominal error bounds. The proposed method systematically exploits the properties of the jackknife detector in the range domain and derives a tight bound of the integrity risk. The proposed method is evaluated in a worldwide simulation with both single and

dual constellations. Results reveal that the proposed method shows higher system availability than the baseline ARAIM method, making it possible to support localizer performance with vertical guidance (LPV) with a decision height of 200 ft using the GPS-Galileo dual constellation.

(5) Performance evaluation with real data in simulation

In the worldwide evaluation of fault detection and ARAIM algorithms, the range projection of clock and orbit error of the code IF combination is simulated based on authentic experimental data instead of relying on empirical models, thereby enhancing the reliability of the experimental results. By adopting this methodology, this thesis exposes the vulnerability of the baseline ARAIM model to support the LPV-200 precision approach under the GPS-Galileo dual-constellation setting. At the same time, the capability of the proposed method to facilitate integrity applications under stringent navigation requirements, particularly at the LPV-200 level, is also revealed.

1.3 Thesis Outline

The remaining part of this thesis is organized as follows: Chapter 2 first gives a brief introduction to the fundamentals of global navigation satellite system (GNSS) positioning, to align notations throughout the thesis. Then the ARAIM architecture and the baseline ARAIM algorithms are described, which provides an overview of the standard procedure in integrity monitoring, including range error modeling, fault detection, integrity risk evaluation, and protection level calculation. Chapter 3 describes the procedure for calculating the orbit and clock errors, as well as defining satellite range error for integrity. Moreover, the nominal performance of SISRE for GPS and Galileo satellites is characterized using historical data, which provides empirical observations for the essence of developing non-Gaussian overbounding methods. Consequently, a sharp yet conservative overbound for heavy-tailed error distributions

is constructed in Chapter 4. The bounding performance of the proposed method on the SISRE of GPS and Galileo satellites is analyzed in detail. Chapter 5 develops the jackknife detector, which provides the theoretical foundation for detecting faulty measurements in linearized pseudorange-based positioning systems under non-Gaussian nominal errors. The detection performance of the proposed method is evaluated in a worldwide simulation in both single and multiple fault settings. Chapter 6 develops an integrity monitoring algorithm by leveraging the jackknife detector, which shows significantly higher availability under stringent navigation requirements when compared to the baseline ARAIM algorithm. Finally, Chapter 7 summarizes this thesis and discusses the future work.

Chapter 2

Fundamentals of Advanced Receiver Autonomous Integrity Monitoring

This chapter introduces the fundamentals of GNSS positioning, including the measurement model, error source modeling, and least-squared-based positioning solution. Specific focus has been put on the dual-frequency pseudorange measurements in the dual-constellation system, which provides the basic notion used in the following chapters. In addition, a brief description of the ARAIM architecture and the baseline ARAIM algorithm is given in this chapter, which presents a standard procedure for implementing integrity monitoring algorithms, laying the groundwork for the subsequent chapters of this thesis.

2.1 GNSS Positioning Fundamentals

GNSS satellites continuously broadcast navigation signals across L band frequencies, enabling users to calculate signal travel time from the satellite to the receiver

and satellite coordinates at any epoch [29]. These navigation signals usually comprise ranging codes, the carrier phase signal, and the navigation data. Since GNSS satellites are equipped with precise onboard atomic clocks [30], the time of signal transmission can be identified. By estimating the time taken between transmission and reception, users can obtain pseudorange measurements from ranging codes and carrier phase measurements from the carrier signal, both of which can be regarded as a measure of the apparent distance between the satellite and the receiver. The navigation data contains all the necessary information to allow users to perform the positioning service, including satellite ephemeris, clock bias parameters, almanac and satellite health status [31]. In the following sections, the GNSS measurement and its error modeling are illustrated.

2.1.1 Measurement models

The GNSS signals contain ranging codes and the carrier phase signal, enabling users to calculate the travel time. Regarding the ranging codes, the receiver determines the travel time ΔT by correlating the received Coarse/Acquisition (C/A) code from the satellite with a replica of this code generated in the receiver. This replica is adjusted in time until the maximum correlation is achieved. The measurement $\varrho = c\Delta T$ is known as the pseudorange, where c is the speed of light. Note that the travel time between transmission and reception is determined as the difference in time measured on two distinct clocks. Due to the synchronization errors between satellite and receiver clocks, the pseudorange measurement does not match the geometric distance between the satellite and the receiver. Moreover, other error sources due to signal propagation through the atmosphere (ionosphere and troposphere), instrumental delays (both in satellite and receiver), multipath and receiver noise can also make the pseudorange measurement deviate from the true geometric distance [32]. With considering all these error terms, the pseudorange measurement $\varrho_{j,f}^i$ regarding the satellite i and the

receiver j at frequency f can be modeled as follows:

$$\varrho_{j,f}^i = r_j^i + c(\tau_j - \tau^i) + Td_j^i + Id_{j,f}^i + K_{\varrho,j,f} - K_{\varrho,f}^i + \varepsilon_{\varrho,j,f}^i, \quad (2.1)$$

where r_j^i is the geometry distance; τ_j and τ^i are receiver and satellite clock offsets from the GNSS time scale, respectively; Td_j^i is the tropospheric delay; $Id_{j,f}^i$ is the frequency-dependent ionospheric delay; $K_{\varrho,j,f}$ and $K_{\varrho,f}^i$ are the frequency-dependent receiver and satellite code instrumental delays, respectively. The ionospheric delay, tropospheric delay, and instrumental delay have consolidated models. Nevertheless, these models still have unmodeled components that can be constrained within a certain magnitude. These errors are represented by the residual error term $\varepsilon_{\varrho,j,f}^i$ in the above equation, mainly including ephemeris and clock error $\varepsilon_{orb\&clk}^i$, residual tropospheric error $\varepsilon_{tropo,j}^i$, residual ionospheric error $\varepsilon_{iono,j,f}^i$, and multipath and code noises $\varepsilon_{\rho,user,j,f}^i$, and can be written by

$$\varepsilon_{\varrho,j,f}^i = \varepsilon_{orb\&clk}^i + \varepsilon_{tropo,j}^i + \varepsilon_{iono,j,f}^i + \varepsilon_{\rho,user,j,f}^i, \quad (2.2)$$

where the residual ionospheric error $\varepsilon_{iono,j,f}^i$ and the multipath and code noises $\varepsilon_{\rho,user,j,f}^i$ are frequency dependent. The impacts of the above error sources on the GNSS signal propagation as well as the empirical models will be further illustrated in Section 2.1.2.

Besides the code, the carrier phase can also be utilized to derive a measure of the distance between the satellite and the receiver. Carrier phase measurements are significantly more accurate compared to code measurements. Nevertheless, the carrier phase measurement is subject to ambiguity by an unknown integer number of wavelengths. This ambiguity changes arbitrarily each time the receiver loses the lock on the signal, leading to abrupt jumps or range discontinuities. The carrier phase measurement $\varphi_{j,f}^i$ regarding the satellite i and the receiver j at frequency f can be modelled as

$$\varphi_{j,f}^i = r_j^i + c(\tau_j - \tau^i) + Td_j^i - Id_{j,f}^i + K_{\varphi,j,f} - K_{\varphi,f}^i + n_{j,f}^i + \varepsilon_{\varphi,j,f}^i, \quad (2.3)$$

where $K_{\varphi,j,f}$ and $K_{\varphi,f}^i$ are instrumental delays related to the carrier phase measurement, $n_{j,f}^i$ is the ambiguity term, and $\varepsilon_{\varphi,j,f}^i$ is the residual error remained in the carrier

phase measurement

$$\varepsilon_{\varphi,j,f}^i = \varepsilon_{orb\&clk}^i + \varepsilon_{tropo,j}^i + \varepsilon_{iono,j,f}^i + \varepsilon_{\varphi,user,j,f}^i. \quad (2.4)$$

Note that the ionospheric term has opposite signs for the code and phase measurements.

Given that the ionospheric delay is related to the carrier frequency f , ARAIM users form the IF linear combination by using the multi-frequency GNSS setup. Before applying the above equations to obtain the code IF combination measurements, a new clock definition that refers to the code IF combination of f_A and f_B frequencies is introduced as follows:

$$c\delta_j = c\tau_j + \frac{f_A^2 K_{\varrho,j,f_A} - f_B^2 K_{\varrho,j,f_B}}{f_A^2 - f_B^2} \quad (2.5a)$$

$$c\delta^i = c\tau^i + \frac{f_A^2 K_{\varrho,j,f_A}^i - f_B^2 K_{\varrho,j,f_B}^i}{f_A^2 - f_B^2}. \quad (2.5b)$$

Indeed, the satellite clocks (broadcast or precise) are referred to as the IF combination of codes in Equation (2.5b). With this new clock definition, Equation (2.1) can be re-written as

$$\varrho_{j,f}^i = r_j^i + c(\delta_j - \delta^i) + Td_j^i + Id_{j,f}^i + \tilde{\alpha}_f(K_{\varrho,j,IF} - K_{\varrho,IF}^i) + \varepsilon_{\varrho,j,f}^i, \quad (2.6)$$

where

$$\tilde{\alpha}_f = \frac{f_A^2 f_B^2}{f^2(f_A^2 - f_B^2)}, \quad (2.7)$$

and

$$K_{\varrho,j,IF} = K_{\varrho,j,f_A} - K_{\varrho,j,f_B} \quad (2.8a)$$

$$K_{\varrho,IF}^i = K_{\varrho,f_A}^i - K_{\varrho,f_B}^i \quad (2.8b)$$

are so-called differential code biases (DCB) of the receiver and the satellite, respectively. By defining

$$\gamma = \frac{f_A^2}{f_B^2}, \quad (2.9)$$

the IF combination of codes can be computed as

$$\begin{aligned}
 \varrho_{j,IF}^i &= \frac{\gamma \varrho_{j,f_A}^i - \varrho_{j,f_B}^i}{\gamma - 1} \\
 &= r_j^i + c(\delta_j - \delta^i) + T d_j^i + \varepsilon_{\varrho,j,IF}^i \\
 &= \sqrt{(p_x^i - u_{j,x})^2 + (p_y^i - u_{j,y})^2 + (p_z^i - u_{j,z})^2} + c(\delta_j - \delta^i) + T d_j^i + \varepsilon_{\varrho,j,IF}^i,
 \end{aligned} \tag{2.10}$$

where $\mathbf{p}^i = [p_x^i, p_y^i, p_z^i]^T$ is the position of the i th satellite, $\mathbf{u}_j = [u_{j,x}, u_{j,y}, u_{j,z}]^T$ is the position of the j th receiver in the Earth-centered Earth-fixed (ECEF) coordinate system, and

$$\varepsilon_{\varrho,j,IF}^i = \varepsilon_{orb\&clk}^i + \varepsilon_{tropo,j}^i + \varepsilon_{\varrho,user,j,IF}^i \tag{2.11a}$$

$$\varepsilon_{\varrho,user,j,IF}^i = \frac{\gamma \varepsilon_{\varrho,user,j,f_A}^i - \varepsilon_{\varrho,user,j,f_B}^i}{\gamma - 1}. \tag{2.11b}$$

Note that the DCBs are canceled in the IF combination. In addition, the receiver code instrumental delays K_{ϱ,j,f_A} and K_{ϱ,j,f_B} will be assimilated into the estimation of the receiver clock. The residual error term $\varepsilon_{\varrho,j,IF}^i$ in (2.10) is usually bounded by a zero-mean Gaussian distribution as

$$\varepsilon_{\varrho,j,IF}^i \sim \mathcal{N}(0, (\sigma_{\varrho,j,IF}^i)^2), \tag{2.12}$$

where the standard deviation is formalized by

$$(\sigma_{\varrho,j,IF}^i)^2 = (\sigma_{orb\&clk}^i)^2 + (\sigma_{tropo,j}^i)^2 + (\sigma_{\varrho,user,j,IF}^i)^2. \tag{2.13}$$

Section 2.1.2 elaborates the source of each term in Equation (2.13).

The formulation of the IF combination of phases are detailed in [29], which is omitted for present in this thesis.

2.1.2 Error modeling of code IF combination measurements

As shown in Equation (2.10), the IF combination of code measurements contains several additional time delays related to the signal propagation or the clocks. Since

these errors inevitably translate into the position domain and significantly affect the accuracy and integrity of positioning solutions [33], it is crucial to understand the underlying mechanisms behind these errors and accurately characterize their profiles. The remaining parts of this section briefly reviews the mechanisms and models of the atmospheric delays, orbit and clock errors, and multipath and code noises, all of which contribute to the range errors of the IF combination of code measurements.

(1) Atmospheric effects

In atmospheric conditions, the density of gases and plasma is known to be non-uniform, leading to spatial and temporal fluctuations in the refractive index [34]. These variations cause electromagnetic rays to traverse extended geometric paths compared to free space due to refractive index gradients. Consequently, electromagnetic signals undergo changes in velocity, both in speed and direction, as they propagate through the atmosphere. From a signal delay perspective, the atmosphere can be segmented into two primary components: the ionosphere, where signal delays are frequency-dependent; and the neutral atmosphere, predominantly the troposphere, acting as a non-dispersive medium [32].

The ionosphere encompasses the atmospheric region extending from approximately 60 km to over 2000 km. The ionospheric refraction of GNSS signals can escalate to a few meters depending on the elevations and solar activities [35]. It is worth noting that the ionosphere is a dispersive medium, meaning that the refraction of GNSS signals is frequency-dependent. This frequency dependence allows dual-frequency users to mitigate the ionospheric effect by over 99.9%.

The troposphere is the atmospheric layer between Earth's surface and an altitude of approximately 60 km. The transmission delay caused by the troposphere is influenced by factors such as the temperature, pressure, humidity as well as the locations of the transmitter and receiver antennas. Unlike the ionosphere, the troposphere is a non-dispersive medium for radio waves at GNSS frequencies. Therefore, it is im-

possible to remove the tropospheric effects with the combination of dual-frequency measurements. Various models have emerged to estimate tropospheric delays based on observational data, emphasizing the characterization of both the dry and wet atmospheric components. While the dry atmosphere can be modeled using ideal gas laws from surface pressure and temperature data, the wet component presents greater challenges due to its unpredictability [36]. In Radio Technical Commission for Aeronautics (RTCA) GPS/Wide Area Augmentation System (WAAS) Minimum Operational Performance Standard (MOPS), a simple tropospheric correction model based on a unified mapping function for both dry and wet components is adopted. The residual tropospheric error bound used by RTCA-MOPS-229D [37] is common to all frequencies and signals and is given by

$$\sigma_{tropo,j}^i = 0.12[\text{m}] \frac{1.001}{\sqrt{0.002001 + \sin^2(\theta_j^i[\text{rad}])}}, \quad (2.14)$$

where θ_j^i is the elevation angle.

(2) Orbit and clock errors

The orbit error and clock error both reflect the difference between the contents of the navigation message and reality [38]. Specifically, the Master Control Station (MCS) determines and transmits ephemeris parameters and clock correction parameters to the satellites for rebroadcast in the navigation message. Since these parameters are computed using a curve-fit to estimate the actual satellite position and clock offset, some residual error remains. The residual orbit error can affect ranging errors on the order of 0.8 m, while the residual clock error results in ranging errors that typically vary from 0.3 ~ 4 m [32]. Typically, the primary factors affecting constellation performance are the residual orbit and clock errors. In ARAIM, these errors are bounded by a zero-mean Gaussian distribution with the 1-sigma bound of $\sigma_{orb\&clk}^i$. Chapters 3 and 4 revisit this problem and demonstrate the benefits of using a non-Gaussian model to bound these errors for integrity applications.

(3) Multipath and code noises

The multipath occurs when a GNSS signal reaches the receiver antenna through different paths, stemming from reflections of surroundings [39]. The reflected signals traverse longer paths compared to the direct signals, which produces delayed and weakened copies that can distort the correlation function in the receiver, eventually leading to errors in position, velocity, and time estimations. Multipath errors vary in magnitude based on the receiver's environment, including factors like satellite elevation angle, receiver signal processing, antenna gain pattern, and signal characteristics. In aviation applications, aircraft structures like fuselage and wings mainly contribute to multipath errors [40]. However, in terrestrial applications, multipath errors typically arise from reflections of buildings, terrain features, or other stationary objects in the vicinity of the receiver. Notably, the natural motion of the aircraft should cause the relative phase of the multipath carrier to change quickly, which will in turn cause the multipath error to be noise-like rather than bias-like, as it is at ground receivers. Therefore, statistical models have been proposed to bound such random errors in aviation applications.

Receiver noise and distortion encompass random errors originated from the receiver hardware, such as antenna and cables. These errors introduce biases into both code and phase measurements. For code measurements, thermal noise jitter and interference are the primary sources of receiver noise and distortion. The 1σ value for noise and resolution errors is typically on the order of a decimeter or less in nominal conditions (without external interference). For phase measurements, the 1σ value is typically on the order of millimeters in nominal conditions. Therefore, receiver noise and distortion are usually negligible compared to errors induced by multipath.

Appendix A.1 lists the multipath and code noise error bound ($\sigma_{\rho, user, j, IF}^i$) for GPS and Galileo airborne receivers.

2.1.3 Positioning solution

To simplify notations, we omit the indices j when presenting the system state, measurement, and measurement error vectors during the following derivation without loss of generality.

Let $\rho_j^i = \varrho_{j,IF}^i - Td_j^i + c\delta_j^i$. Then, the measurement model in (2.10) can be linearized by taking the first-order Taylor expansion at $\mathbf{x}_0 = [u_{j,x,0}, u_{j,y,0}, u_{j,z,0}, \delta_{j,0}]^T$ as follows:

$$\begin{aligned} \rho_j^i = & \rho_{j,0}^i - a_{j,1}^i (u_{j,x} - u_{j,x,0}) - a_{j,2}^i (u_{j,y} - u_{j,y,0}) \\ & - a_{j,3}^i (u_{j,z} - u_{j,z,0}) + c(\delta_j - \delta_{j,0}) + \varepsilon_{\varrho,j,IF}^i, \end{aligned} \quad (2.15)$$

where

$$\rho_{j,0}^i = \sqrt{(p_x^i - u_{j,x,0})^2 + (p_y^i - u_{j,y,0})^2 + (p_z^i - u_{j,z,0})^2} + c\delta_{j,0} \quad (2.16a)$$

$$a_{j,1}^i = \frac{p_x^i - u_{j,x,0}}{\sqrt{(p_x^i - u_{j,x,0})^2 + (p_y^i - u_{j,y,0})^2 + (p_z^i - u_{j,z,0})^2}} \quad (2.16b)$$

$$a_{j,2}^i = \frac{p_y^i - u_{j,y,0}}{\sqrt{(p_x^i - u_{j,x,0})^2 + (p_y^i - u_{j,y,0})^2 + (p_z^i - u_{j,z,0})^2}} \quad (2.16c)$$

$$a_{j,3}^i = \frac{p_z^i - u_{j,z,0}}{\sqrt{(p_x^i - u_{j,x,0})^2 + (p_y^i - u_{j,y,0})^2 + (p_z^i - u_{j,z,0})^2}}. \quad (2.16d)$$

Assuming there are n measurements from N_{const} constellations, we can redefine the receiver true state as

$$\mathbf{x}_t = [u_{j,x}, u_{j,y}, u_{j,z}, -\delta_{j,1}, \dots, -\delta_{j,N_{const}}]^T, \quad (2.17)$$

where $\delta_{j,c}, c = 1, 2, \dots, N_{const}$ is the receiver clock offset corresponding to the c th constellation. For notations, define

$$m = 3 + N_{const} \quad (2.18)$$

as the size of the state. Similarly, the linearization point is defined as

$$\mathbf{x}_0 = [u_{j,x,0}, u_{j,y,0}, u_{j,z,0}, -\delta_{j,1,0}, \dots, -\delta_{j,N_{const},0}]^T. \quad (2.19)$$

The matrix form of the linearized measurement model with n measurements from N_{const} constellations can be written as

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \boldsymbol{\varepsilon}, \quad (2.20)$$

where

$$\mathbf{y} = \begin{bmatrix} \rho_{j,0}^1 - \rho_j^1 \\ \vdots \\ \rho_{j,0}^n - \rho_j^n \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{\ell,j,IF}^1 \\ \vdots \\ \varepsilon_{\ell,j,IF}^n \end{bmatrix}, \mathbf{x} = \mathbf{x}_t - \mathbf{x}_0, \quad (2.21)$$

and the elements of the $n \times m$ matrix $\mathbf{G}$ are defined by

- $G_{i,1} = a_{j,1}^i, G_{i,2} = a_{j,2}^i, G_{i,3} = a_{j,3}^i$;
- $G_{i,3+c} = 1$ if satellite i belongs to constellation q ;
- $G_{i,3+c} = 0$ if satellite i does not belong to constellation q .

The system state $\mathbf{x}$ can be solved by the weighted least square (WLS) method (in an iterative approach) as follows:

$$\hat{\mathbf{x}} = \mathbf{S}\mathbf{y}, \quad (2.22)$$

where $\mathbf{S}$ is the solution matrix

$$\mathbf{S} = (\mathbf{G}^T \mathbf{W} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}, \quad (2.23)$$

and $\mathbf{W}$ is the weight matrix and will be defined in Section 2.3.2.

2.2 ARAIM Architecture

GNSS measurements are vulnerable to infrequent faults such as satellite and constellation failures, which have the potential to pose significant safety risks to users. To address this issue, fault detection and exclusion (FDE) algorithms such as legacy RAIM can be implemented. The legacy RAIM is designed to provide error bounds

for lateral navigation but faces limitations in precise vertical navigation due to factors like ionospheric model uncertainty and high DOP.

The modernization of GPS [11–13] and the emerging GNSS constellations present the chance to elevate legacy RAIM to ARAIM. ARAIM builds upon the foundation of legacy RAIM by incorporating multi-frequency measurements, which allow for the utilization of IF linear combinations to remove ionospheric uncertainties and reduce radio frequency interference. ARAIM also enhances the strength of satellite geometries to provide less conservative protection levels, and introduces the constellation fault monitors, addressing the limitations of the legacy RAIM. These advancements in ARAIM aim to support more robust and precise aircraft guidance in civil aviation.

The ARAIM system architecture consists of the space segment, ground segment, airborne segment, and Integrity Support Message [41]. Specifically, the space segment encompasses the GNSS core constellations managed by their respective constellation service providers (CSP) and involves various operations like monitoring stations, mission segments, orbit determination and time synchronization (ODTS), and performance commitments. The ground segment involves reference stations responsible for monitoring constellation performance and collecting data for integrity support message (ISM) generation. The airborne segment mainly comprises aviation users equipped with ARAIM avionics. The ISM delivers essential inputs for users to assess performance metrics, such as the prior probability of satellite and constellation faults, the standard deviation of nominal ranging uncertainty due to satellite orbit and clock errors, and the maximum nominal bias primarily due to signal deformation. The Worldwide GNSS Committee (WGC) recognizes two services, Horizontal ARAIM and Vertical ARAIM, for different flight phases. Vertical ARAIM is designed to support global LPV-200 precision approach operations [27] through either offline or online ARAIM architectures. Further details on Horizontal ARAIM, offline and online ARAIM are elaborated in [41]. Subsequent sections will delve into the ARAIM user algorithms in the airborne segment.

2.3 ARAIM Airborne Algorithm

2.3.1 Concepts of integrity monitoring

Integrity comprises several factors such as integrity risk, alert limit (AL), and time to alert (TTA), representing the reliability of a navigation solution in critical situations. The boundary defining unsafe errors for a specific application is referred to as the AL, whereas the hazardous condition we strive to prevent is termed misleading information (MI). Loss of integrity arises when MI exists without notification or mitigation within the required TTA. Mathematically, integrity risk (P_{HMI}) can be defined as

$$P_{\text{HMI}} = P(|e_0| > \ell \cap q \notin R) \leq I_{\text{REQ}}, \quad (2.24)$$

where e_0 is the estimation error on the parameter of interest (e.g., the vertical positioning solution), ℓ is the specified AL, q and R are a general test statistic and its rejection region, respectively, and I_{REQ} is the integrity budget. The design of test statistics and the choice of rejection region distinguish different integrity monitoring algorithms [2, 18, 20].

Another important concept in integrity monitoring is continuity risk, which quantifies the probability of unforeseen navigation loss necessitating the termination of operations to ensure safety. Three primary factors, including satellite faults, malfunctioning GNSS system components, and false alerts from integrity monitors, contribute to continuity loss in GNSS applications. Notably, the determination of the rejection region for integrity monitors is influenced by the false alert as follows:

$$P(q \in R | \text{Fault-free conditions}) P_{\text{Fault-free}} \leq C_{\text{REQ,FA}}, \quad (2.25)$$

where $C_{\text{REQ,FA}}$ is the continuity budget caused by false alerts, and $P_{\text{Fault-free}}$ is the prior probability of no fault, which is close to 1.

As GNSS users usually have limited control over mitigating satellite faults and failures in system components compared to managing the false alert rate, integrity monitor-

ing algorithms typically concentrate on minimizing false alert risk, leading to a direct trade-off between the continuity and integrity requirements. Specifically, setting the rejection region of integrity monitors to its maximum allows for detecting and removing faulty measurements within TTA to prevent integrity loss. However, widening the rejection region increases the likelihood of excluding fault-free measurements due to random noise under normal conditions. Balancing this trade-off to meet both integrity and continuity requirements simultaneously is a key challenge for designers of integrity monitoring algorithms with stringent safety-of-life demands. The following sections introduce the core processes of the baseline ARAIM airborne algorithm [20], covering range error bound, fault detection, and protection level calculations.

2.3.2 Range error bound

The baseline ARAIM airborne algorithm employs two distinct error models, i.e., the error model for accuracy and the error model for integrity [20], as listed below:

(1) Error model for accuracy and continuity

The range error is bounded by a zero-mean Gaussian distribution as follows:

$$\varepsilon_{\rho,j,IF}^i \sim \mathcal{N}(0, (\sigma_{acc,j}^i)^2) , \quad (2.26)$$

where

$$(\sigma_{acc,j}^i)^2 = (\sigma_{URE}^i)^2 + (\sigma_{tropo,j}^i)^2 + (\sigma_{\rho,user,j,IF}^i)^2 , \quad (2.27)$$

and σ_{URE}^i is the 1-sigma orbit and clock error bound. Notably, the nominal bias for continuity is assumed to be zero. The covariance matrix for accuracy is then defined as

$$\mathbf{C}_{acc}(i, i) = (\sigma_{acc,j}^i)^2 . \quad (2.28)$$

(2) Error model for integrity

The range error is bounded by a Gaussian distribution as follows

$$\varepsilon_{\rho,j,IF}^i \sim \mathcal{N}(b_{nom}^i, (\sigma_{int,j}^i)^2) , \quad (2.29)$$

where

$$(\sigma_{int,j}^i)^2 = (\sigma_{URA}^i)^2 + (\sigma_{tropo,j}^i)^2 + (\sigma_{\rho,user,j,IF}^i)^2, \quad (2.30)$$

σ_{URA}^i (the broadcast user range accuracy (URA) parameter) is the 1-sigma orbit and clock error bound, and b_{nom}^i is the nominal bias for integrity. The covariance matrix for integrity is defined as

$$\mathbf{C}_{int}(i, i) = (\sigma_{int,j}^i)^2. \quad (2.31)$$

2.3.3 Integrity and continuity under multiple hypothesis

The baseline ARAIM airborne algorithm employs a multiple-hypothesis approach by first defining the threat model, which is a collection of error modes that partition the whole measurement space. Assuming there are n measurements each uniquely numbered, the threat model is constructed by defining a set of fault modes with different prior probabilities:

- Fault mode 0: All measurements are nominal measurements (i.e., fault-free). The prior probability of fault mode 0 is P_{H_0} .
- Fault mode k : Measurements with indices $k \in idx_k^{ex}$ are faulty measurements (including single or multiple faults), while measurements with indices $k \notin idx_k^{ex}$ are nominal measurements. The prior probability of fault mode k is P_{H_k} .

In the above definition, the size of idx_k^{ex} is the number of simultaneous faults associated with the fault mode k , which takes value from 1 to n . The total number of fault modes is assumed to be $N_{\text{fault modes}} + 1$. Theoretically,

$$N_{\text{fault modes}} = \sum_{k=1}^{n-k_{\max}} \binom{n}{k}, \quad (2.32)$$

where $k_{\max}$ is the maximum number of simultaneous faults that need to be monitored. $k_{\max}$ is selected so that the prior probability of occurrence of more than $k_{\max}$ simultaneous faults is much smaller than I_{REQ} . This probability is denoted as $P_{\text{not monitored}}$.

The procedure for determining $k_{\max}$ and P_{H_i} is detailed in [20]. For information, this procedure is also briefed in Appendix A.2.

Figure 2.1 conceptualizes the definition of fault modes, where blue dots represent fault-free measurements and dashed circles indicate faulty measurements under each fault mode. Then, the hypotheses in the baseline ARAIM airborne algorithm are given by

- H_0 : The fault-free hypothesis (fault mode 0).
- H_k : The hypothesis corresponding to fault mode k .

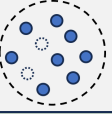
Fault mode	H_0	H_1	H_2	H_k
Meas. subset				
Solution	$\hat{x}$	$\hat{x}^{(1)}$	$\hat{x}^{(2)}$	$\hat{x}^{(k)}$

Figure 2.1: Illustration of the multiple hypothesis solution separation in the baseline ARAIM airborne algorithm.

The integrity risk in Equation (2.24) can be rewritten by

$$P_{\text{HMI}} = \sum_{i=0}^{N_{\text{fault modes}}} P \left(\{|e_0| > \ell\} \cap \bigcap_{k=1}^{N_{\text{fault modes}}} q_k \notin R_k \mid H_k \right) P_{H_i} + P_{\text{not monitored}} \leq I_{\text{REQ}}, \quad (2.33)$$

where q_k and R_k are a general test statistic and its rejection region regarding the test H_0 v.s. H_k , respectively.

The continuity risk is given by

$$P_{\text{continuity}} = P \left(\bigcup_{k=1}^{N_{\text{fault modes}}} q_k \in R_k \mid H_0 \right) P_{H_0} \leq C_{\text{REQ,FA}}. \quad (2.34)$$

2.3.4 Solution separation based fault detection

The baseline ARAIM adopts the solution separation (SS) between different measurement subsets as the test statistic. The full set solution associated with fault mode 0 is given in Equation (2.22) and can be re-written by

$$\hat{\mathbf{x}}^{(0)} = \mathbf{S}^{(0)} \mathbf{y}, \quad (2.35)$$

where

$$\mathbf{S}^{(0)} = (\mathbf{G}^T \mathbf{W}^{(0)} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}^{(0)}, \quad (2.36)$$

and $\mathbf{W}^{(0)}$ is associated with the integrity range error bound as follows

$$\mathbf{W}^{(0)} = \mathbf{C}_{int}^{-1}. \quad (2.37)$$

For fault mode k , the measurements with indices $k \notin idx_k^{ex}$ are used to compute the subset solution through WLS

$$\hat{\mathbf{x}}^{(k)} = \mathbf{S}^{(k)} \mathbf{y}, \quad (2.38)$$

where

$$\mathbf{S}^{(k)} = (\mathbf{G}^T \mathbf{W}^{(k)} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}^{(k)} \quad (2.39)$$

is the solution matrix on fault mode k , and $\mathbf{W}^{(k)}$ is defined as [20]

$$\mathbf{W}^{(k)}(i, i) = \begin{cases} \mathbf{C}_{int}^{-1}(i, i) & \text{if } i \notin idx_k^{ex} \\ 0 & \text{otherwise} \end{cases}. \quad (2.40)$$

The variance of the positioning solution is given by

$$(\sigma_v^{(k)})^2 = (\mathbf{G}^T \mathbf{W}^{(k)} \mathbf{G})_{v,v}^{-1}, \quad (2.41)$$

where the subscript $v = 1, 2, 3$ designates the east, north, and up components, respectively. Notably, the nominal biases b_{nom}^i of the integrity range error bound also have impacts on the positioning solution. The worst-case impact on the position solution $\hat{\mathbf{x}}^{(k)}$ is given by

$$b_v^{(k)} = \sum_{i=1}^n |\mathbf{S}_{v,i}^{(k)}| b_{nom}^i. \quad (2.42)$$

The test statistic is constructed by the separation between the full solution and the k th subsolution as

$$d_v^{(k)} = (\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(k)})_v. \quad (2.43)$$

Since the residual error $\varepsilon_{\ell,j,IF}^i$ is assumed to have a Gaussian distribution, $d_v^{(k)}$ is proven to have the following Gaussian distribution under the fault-free hypothesis

$$d_v^{(k)} \sim \mathcal{N}\left(0, (\sigma_{ss,v}^{(k)})^2\right), \quad (2.44)$$

where

$$(\sigma_{ss,v}^{(k)})^2 = \mathbf{e}_v^T (\mathbf{S}^{(k)} - \mathbf{S}) \mathbf{C}_{acc} (\mathbf{S}^{(k)} - \mathbf{S})^T \mathbf{e}_v, \quad (2.45)$$

and $\mathbf{e}_v$ is a vector whose v th entry is 1 and all others are 0. Then the rejection region is given by

$$R_k = \{d_v^{(k)} \mid |d_v^{(k)}| > D_{k,v}\}, \quad (2.46)$$

where $D_{k,v}$ is the threshold of the integrity monitor. Therefore, the continuity risk in (2.34) can be written by

$$P_{\text{continuity}} = P\left(\bigcup_{k=1}^{N_{\text{fault modes}}} |d_v^{(k)}| > D_{k,v} \mid H_0\right) P_{H_0}. \quad (2.47)$$

Since these rejection regions $R_k, k = 1, 2, \dots, N_{\text{fault modes}}$ are not mutually exclusive, Equation (2.47) can be released by

$$P_{\text{continuity}} \leq \sum_{k=1}^{N_{\text{fault modes}}} P\left(|d_v^{(k)}| > D_{k,v} \mid H_0\right) P_{H_0}. \quad (2.48)$$

The threshold $D_{k,v}$ is determined by the allocated continuity budget caused by false alerts. Specifically,

$$D_{k,v} = Q^{-1}\left(\frac{C_{\text{REQ,FA},j}^v}{2P_{H_0}}\right) \sigma_{ss,v}^{(k)}, \quad (2.49)$$

where

$$C_{\text{REQ,FA}}^v = \sum_{j=1}^{N_{\text{fault modes}}} C_{\text{REQ,FA},j}^v, \quad (2.50)$$

and $Q^{-1}(\cdot)$ is the quantile function of the standard normal distribution. Here, $C_{\text{REQ,FA}}^3$ standards for the continuity budget for the vertical component, $C_{\text{REQ,FA}}^1 + C_{\text{REQ,FA}}^2$ represents the continuity budget for the horizontal component, and $C_{\text{REQ,FA}}^1 = C_{\text{REQ,FA}}^2$. An equal allocation strategy of the continuity budget is adopted in the baseline ARAIM algorithm, i.e.,

$$C_{\text{REQ,FA},j}^v = \frac{1}{N_{\text{fault modes}}} C_{\text{REQ,FA}}^v. \quad (2.51)$$

A potential fault is detected when $d_v^{(k)} \in R_k, \forall v = 1, 2, 3$, and then an exclusion process is launched, which is beyond the scope of this thesis and is omitted for the present. The detailed process of exclusion and its impacts on the system integrity can be referred to [20]. The following section will illustrate the evaluation of system integrity in nominal conditions where no faults are detected.

2.3.5 Calculation of protection levels

Instead of directly evaluating the integrity risk, the baseline ARAIM algorithm calculates the PL, the position error bounds at the low probabilities necessary to meet the integrity requirement for a specified application. Practically, integrity is maintained as long as the PL calculated at each epoch remains below the AL. By replacing the alert limit ℓ with PL and replacing the last inequality with equality in Equation (2.33), the PL in the baseline ARAIM algorithm is given by

$$\begin{aligned} & 2Q\left(\frac{PL_v - b_3^{(0)}}{\sigma_3^{(0)}}\right) + \sum_{k=1}^{N_{\text{fault modes}}} P_{H_k} Q\left(\frac{PL_v - D_{k,3} - b_3^{(k)}}{\sigma_3^{(k)}}\right) \\ & = I_{\text{REQ}}^v \left(1 - \frac{P_{\text{not monitored}}}{I_{\text{REQ}}}\right), \end{aligned} \quad (2.52)$$

where I_{REQ}^3 standards for the integrity budget for the vertical component, $I_{\text{REQ}}^1 + I_{\text{REQ}}^2$ represents the integrity budget for the horizontal component, and $I_{\text{REQ}}^1 = I_{\text{REQ}}^2$. The vertical protection level (VPL) is directly given by

$$VPL = PL_3, \quad (2.53)$$

and the horizontal protection level (HPL) is calculated by

$$HPL = \sqrt{PL_1^2 + PL_2^2}. \quad (2.54)$$

Notably, Equation (2.52) involves the detection threshold $D_{k,v}$, which is determined by the allocated continuity budget caused by false alerts.

The baseline ARAIM algorithm provides a standard procedure, including range error modeling, fault detection, and protection level calculation, to support integrity applications. The whole algorithm is built on the Gaussian overbounds for nominal range errors, which simplifies the derivation and reduces the computational effort. However, nominal range errors in the real world usually have non-Gaussian and heavy-tailed properties [22–24, 42, 43]. The Gaussian overbounds of these errors are likely to be over-conservative, leading to degraded detection performance and inflated PLs in real-world applications. To solve this issue, this thesis systematically incorporates non-Gaussian modeling into the construction of the integrity monitoring algorithm, including characterization of the range projection of orbit and clock errors (Chapter 3), error bounding with non-Gaussian models (Chapter 4), and fault detection (Chapter 5) and integrity risk evaluation (Chapter 6) with non-Gaussian nominal bounds.

Chapter 3

Characterization of GPS/Galileo Signal-In-Space Range Error

Navigation satellites transmit electromagnetic waves known as signal-in-space (SIS) to users and GNSS ground segments [44]. The performance of SIS is commonly assessed by analyzing the SISRE, which quantifies the statistical uncertainty of the modeled pseudorange due to errors in the broadcast orbit and clock information [45,46]. SISRE is one of the most significant factors that degrade positioning accuracy and integrity [47]. Therefore, understanding the individual SISRE contributions and characterizing their statistical profiles is essential for the development of ARAIM algorithms.

This chapter aims to characterize GPS/Galileo SISRE to facilitate the development of non-Gaussian overbounding and ARAIM methods in Chapter 4 and Chapter 6, respectively.

3.1 Sources of Signal-In-Space Range Error

SISRE describes the statistical uncertainty of the modeled pseudorange due to errors in the broadcast orbit and clock information [25, 45, 46]. Satellite orbit and clock errors arise due to uncertainties in the ODTS process managed by the CSP [25]. Typically, the CSP utilizes dynamic models to establish a reference orbit trajectory. Sequential estimators, such as Kalman filters, are then deployed to refine this trajectory, forecasting the future evolution of satellite orbit and clock [32, 48]. These predictions are then employed to determine broadcast ephemeris parameters, which are uploaded into the satellite's navigation payload memory and transmitted to the GNSS user. The estimation error of reference orbit trajectory and the correction error of sequential estimators together contribute to the broadcast ephemeris errors.

A typical example is given by the ODTS process employed in the MCS of GPS. A dynamic model is constructed by integrating a gravity model, luni-solar perturbations, solid Earth tides, and a priori solar radiation pressure model. The propagation of a satellite orbit is then solved by numerical integration of the dynamics model for given initial values. Meanwhile, the evolution of clock states is characterized by second-order polynomials over time. Monitor stations collect raw pseudorange, carrier phase, and meteorological data, which are smoothed by the MCS and fed into a Kalman filter for processing. This filter estimates corrections of the inertial satellite positions and velocities relative to the reference trajectory, two radiation pressure parameters, and clock offset, drift, and rates for both satellites and monitoring stations. These estimations are subsequently utilized to predict the future orbit and clock evolution, and to modify the broadcast ephemeris parameters for the next upload to the satellite. A similar procedure is applied for other constellations such as GLObalnaya NAVigatsionnaya Sputnikovaya Sistema in Russian (GLONASS) and Galileo.

3.2 Determination of Orbit and Clock Errors

The broadcast orbit and clock errors can be evaluated as discrepancies between the satellite position and clock bias given in the broadcast ephemeris (BCE) and the precise ephemeris (PCE). The data source of the BCEs and PCEs used in this study is elaborated in Section 3.2.1. Because the BCEs and PCEs are referred to the different time and coordinate systems, adequate care must be taken to ensure consistency in the comparison. Aside from time and coordinate system alignment, several additional corrections are required in the data preprocessing procedures, which will be illustrated in Section 3.2.3. Thereafter, the computation method of orbit and clock errors is given.

3.2.1 Data source

This study is carried out by employing the historical broadcast and precise ephemerides of GPS and Galileo. The performance evaluation for GPS is conducted with respect to L1/L2 combination over a three-year period from January 1st, 2020 to December 31st, 2022. The analysis for Galileo satellites is conducted with respect to E1/E5a combination within the same period. Regarding decommissioned satellites, this study excludes satellites that were retired from duty before December 2022. For GPS satellites, SVN 34 and SVN 60 were retired from active duty in October 2019 and March 2020, respectively [49]. Without considering these two satellites, a total number of 30 GPS satellites are used for analysis in this study. For Galileo satellites, GSAT0204 was decommissioned from active service for constellation management in December 2017 [50], and GSAT0201 and GSAT0202 became unavailable from February 18th, 2021 due to a temporary setting of the Freely Accessible Navigation Message (F/NAV) and Integrity Navigation Message (I/NAV) health bits to unhealthy [50]. Therefore, these three satellites are not included in this study. A total number of 24 Galileo satellites are utilized for further analysis in this three-year period.

The BCEs for GPS and Galileo satellites are both acquired from International GNSS Service (IGS), which records multi-GNSS navigation data on a daily basis and compiles them into the BRDC file in Receiver Independent Exchange (RINEX) format Version 3 [51]. Various IGS and Multi-GNSS EXperiment (MGEX) [52] analysis centers routinely generate PCEs for GPS and GLObalnaya NAvigatsionnaya Sputnikovaya Sistema in Russian (GLONASS), along with the emerging Galileo, BeiDou, and Quasi-Zenith Satellite System (QZSS) constellations. In this study, the precise orbit and clock products of GPS are obtained from Center for Orbit Determination in Europe (CODE), which provides daily orbit and clock data with 15-minute sampling. The precise orbit and clock products of Galileo are also obtained from CODE but with the sampling interval of 5 minutes. The precise products of CODE are publicly accessible through the online repository of IGS [53].

3.2.2 Antenna phase center offset correction

According to IGS conventions, most precise orbit products provide the Center of Mass (CoM) coordinates of the respective space vehicle [54]. However, the precise clock states are usually referenced to satellite antenna phase center (APC) [44]. In the case of broadcast messages, both orbit and clock data are referred to the satellite APC [45]. Figure 3.1 depicts the orbit and clock reference points in broadcast and precise ephemerides. Due to the different conventions, a proper APC offset correction is needed before comparing the broadcast and precise ephemerides. However, as pointed out by [45, 55], there is no reason to assume a uniquely and unambiguously defined APC between the broadcast ephemeris and precise orbit and clock products. In this work, the estimated APC offsets by the US National Geospatial-Intelligence Agency (NGA) [56] are used for the GPS BCE, while the Galileo metadata provided by the European GNSS Agency (GSA) [57] is used to generate the APC offsets for the Galileo BCE. In the case of precise products, both the Galileo and GPS APC offsets are provided by the IGS ANTenna EXchange (ANTEX) files [58].

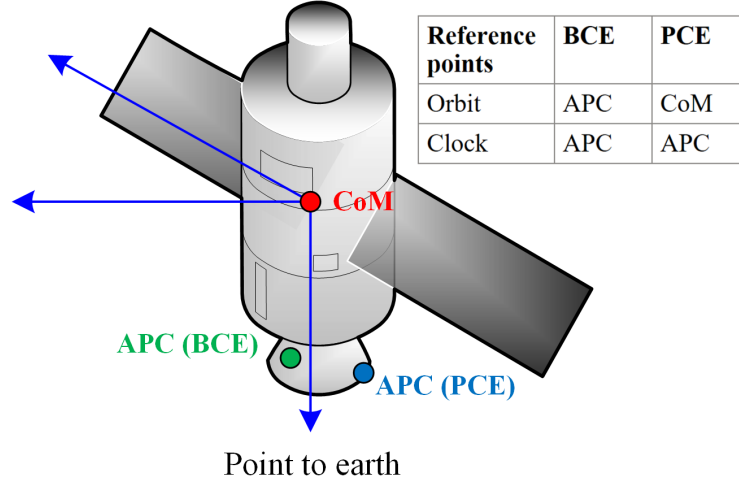


Figure 3.1: Different conventions for orbit and clock reference points in GNSS broadcast and precise ephemerides. The green dot represents the APC used in BCE and the blue dot represents the APC used in PCE.

The broadcast CoM position $\mathbf{x}_{CoM,ECEF}^{i,k,BCE}$ of satellite i at epoch k in the ECEF coordinate system is obtained by applying the APC offset correction as follows:

$$\mathbf{x}_{CoM,ECEF}^{i,k,BCE} = \mathbf{x}_{APC,ECEF}^{i,k,BCE} - \mathbf{R}_{BF}^{i,k,ECEF} \mathbf{a}_{BCE}^i, \quad (3.1)$$

where $\mathbf{x}_{APC,ECEF}^{i,k,BCE}$ is the broadcast APC position, $\mathbf{a}_{BCE}^i = [a_{x,BCE}^i, a_{y,BCE}^i, a_{z,BCE}^i]^T$ is the APC offset vector for the BCE defined in the satellite body fixed (BF) frame, and $\mathbf{R}_{BF}^{i,k,ECEF}$ is the rotation matrix from the BF frame to the ECEF frame at epoch k .

Figure 3.2 shows the definition of the BF frame, where $\mathbf{e}_{z,BF}$ is the principal body axis closest to the antenna boresight direction, $\mathbf{e}_{y,BF}$ is parallel to the rotation axis of the solar panels, and $\mathbf{e}_{x,BF}$ is chosen such that the specified solar panel is permanently sunlit during nominal yaw-steering in IGS conventions [59]. Define the satellite-sun vector as

$$\mathbf{x}_{sat \text{ to sun}}^{i,k} = \mathbf{x}_{sun}^k - \mathbf{p}^{i,k}, \quad (3.2)$$

where $\mathbf{p}^{i,k}$ and $\mathbf{x}_{sun}^k$ are the position of satellite i and the sun at epoch k in the ECEF frame, respectively. Then, the satellite-sun vector $\mathbf{x}_{sat \text{ to sun}}^{i,k}$ is orthogonal to the solar

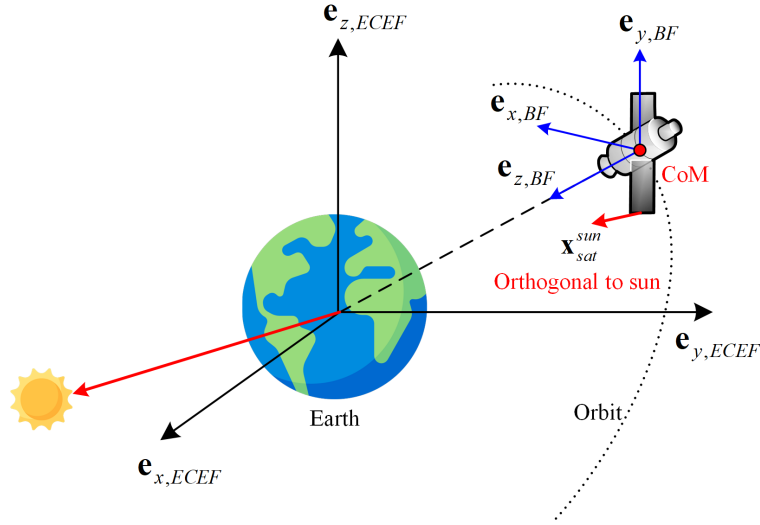


Figure 3.2: Definition of the BF frame of satellites.

panel, as shown in Figure 3.2. The rotation matrix $\mathbf{R}_{BF}^{i,k,ECEF}$ can be calculated by

$$\mathbf{R}_{BF}^{i,k,ECEF} = \begin{bmatrix} \frac{\mathbf{x}_{sat}^{i,k} \text{ to sun} \times \mathbf{p}^{i,k} \times \mathbf{p}^{i,k}}{|\mathbf{x}_{sat}^{i,k} \text{ to sun} \times \mathbf{p}^{i,k}| \cdot |\mathbf{p}^{i,k}|} & -\frac{\mathbf{x}_{sat}^{i,k} \text{ to sun} \times \mathbf{p}^{i,k}}{|\mathbf{x}_{sat}^{i,k} \text{ to sun} \times \mathbf{p}^{i,k}|} & -\frac{\mathbf{p}^{i,k}}{|\mathbf{p}^{i,k}|} \end{bmatrix}. \quad (3.3)$$

Since the line-of-sight (LOS) vector is closely aligned with the $\mathbf{e}_{z,BF}$ axis, the z-component of the APC offset $a_{z,BCE}^i$ can be used to calculate the CoM-referenced clock offset $c\delta_{CoM,BCE}^{i,k}$ as follows,

$$c\delta_{CoM,BCE}^{i,k} = c\delta_{APC,BCE}^{i,k} + a_{z,BCE}^i, \quad (3.4)$$

where $c\delta_{APC,BCE}^{i,k}$ is the APC-referenced clock offset in the BCE.

The CoM position of precise orbit $\mathbf{x}_{CoM,ECEF}^{i,k,PCE}$ can be directly computed from the precise product without APC offset corrections. The CoM-referenced clock offset $c\delta_{CoM,BCE}^{i,k}$ in the precise product is computed by

$$c\delta_{CoM,PCE}^{i,k} = c\delta_{APC,PCE}^{i,k} + a_{z,PCE}^i, \quad (3.5)$$

where $c\delta_{APC,BCE}^{i,k}$ is the APC-referenced clock offset in the PCE and $a_{z,PCE}^i$ is the z-component of the APC offset vector $\mathbf{a}_{PCE}^i$ related to the precise product.

In the following sections, the position and clock offset values obtained from broadcast and precise products are consistently referred to the CoM.

3.2.3 Computation of orbit and clock error vector

As shown in Figure 3.3, the orbit error vector of satellite i at time epoch k in the ECEF coordinate system can be formalized by

$$\xi_{orb,ECEF}^{i,k} = \mathbf{x}_{CoM,ECEF}^{i,k,BCE} - \mathbf{x}_{CoM,ECEF}^{i,k,PCE}. \quad (3.6)$$

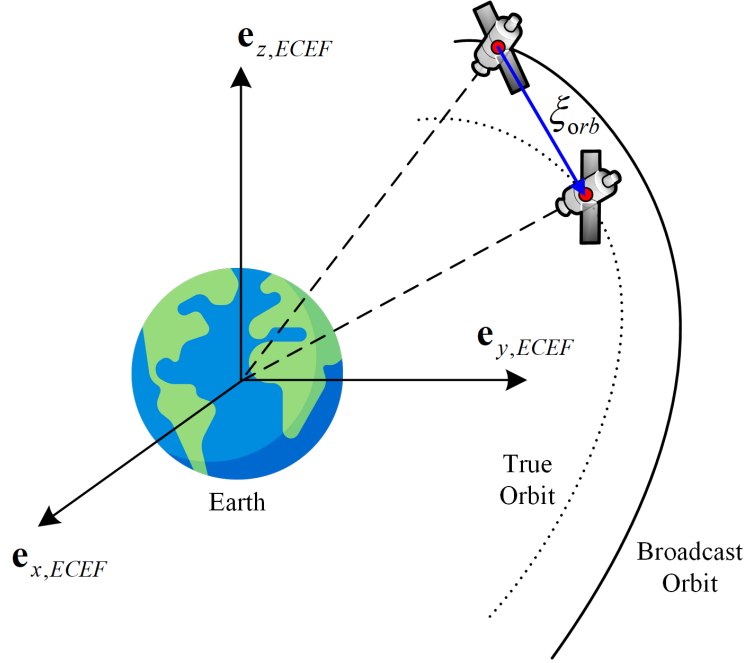


Figure 3.3: Demonstration of orbit error.

The CoM-referenced clock error is computed as

$$\xi_{clk,ECEF}^{i,k} = c\delta_{CoM,BCE}^{i,k} - c\delta_{CoM,PCE}^{i,k}. \quad (3.7)$$

Notably, the underlying realization of the GNSS-specific system time scales is different between broadcast and precise ephemerides [45]. These differences, which are typically more significant than the inherent clock solution precision, introduce a systematic bias

affecting all satellites within a constellation consistently. [55] This bias may vary from epoch to epoch and cannot be easily quantified. To mitigate this bias, an ensemble clock difference is computed at each epoch by averaging the clock error between broadcast and precise ephemerides across satellites within a constellation [44, 55]. Finally, the individual clock error is corrected for this ensemble average as follows:

$$\tilde{\xi}_{clk,ECEF}^{i,k} = \xi_{clk,ECEF}^{i,k} - \frac{1}{N_{sat,k}} \sum_{i=1}^{N_{sat,k}} \xi_{clk,ECEF}^{i,k}, \quad (3.8)$$

where $N_{sat,k}$ is the total number of satellites in the corresponding constellation at epoch k .

3.2.4 Transformation to the RAC frame

The orbit error in the ECEF frame is usually transformed into the radial, along-track and cross-track (RAC) frame for ease of analysis and visualization, as shown in Figure 3.4. The transformation can be formalized by

$$\boldsymbol{\xi}_{orb,RAC}^{i,k} = \mathbf{R}_{ECEF}^{i,k,RAC} \boldsymbol{\xi}_{orb,ECEF}^{i,k} \quad (3.9)$$

with the rotation matrix $\mathbf{R}_{ECEF}^{i,k,RAC}$ defined as

$$\mathbf{R}_{ECEF}^{i,k,RAC} = \begin{bmatrix} -\frac{\mathbf{p}^{i,k}}{|\mathbf{p}^{i,k}|} & -\frac{\mathbf{p}^{i,k} \times \mathbf{v}^{i,k}}{|\mathbf{p}^{i,k} \times \mathbf{v}^{i,k}|} \times \frac{\mathbf{p}^{i,k}}{|\mathbf{p}^{i,k}|} & \frac{\mathbf{p}^{i,k} \times \mathbf{v}^{i,k}}{|\mathbf{p}^{i,k} \times \mathbf{v}^{i,k}|} \end{bmatrix}^T, \quad (3.10)$$

where $\mathbf{v}^{i,k}$ is the velocity of satellite i at epoch k in the ECEF frame. The radial, along-track and cross-track components of $\boldsymbol{\xi}_{orb,RAC}^{i,k}$ are denoted by $\xi_{orb,R}^{i,k}$, $\xi_{orb,A}^{i,k}$, and $\xi_{orb,C}^{i,k}$, respectively.

3.3 Definition of SISRE

To reveal the effect of orbit and clock errors on a navigation user, SISRE is calculated by projecting the orbit and clock errors into the user's LOS direction. Since

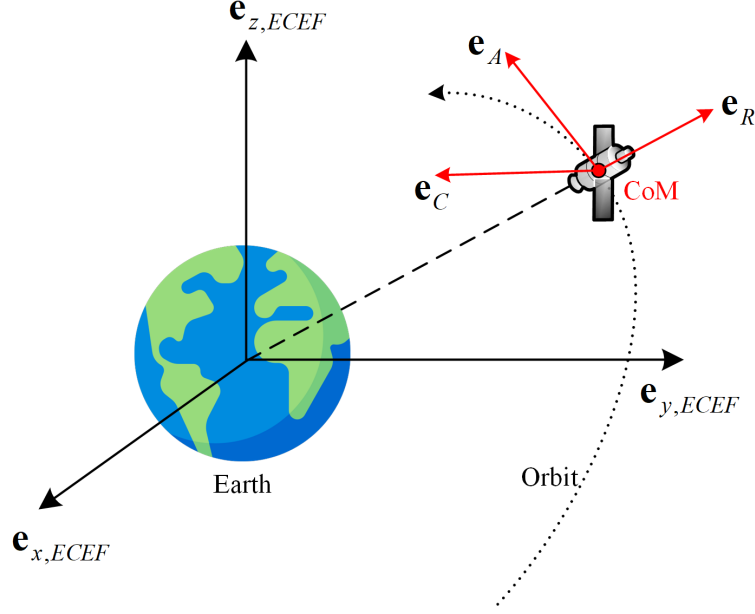


Figure 3.4: Definition of satellite RAC frame.

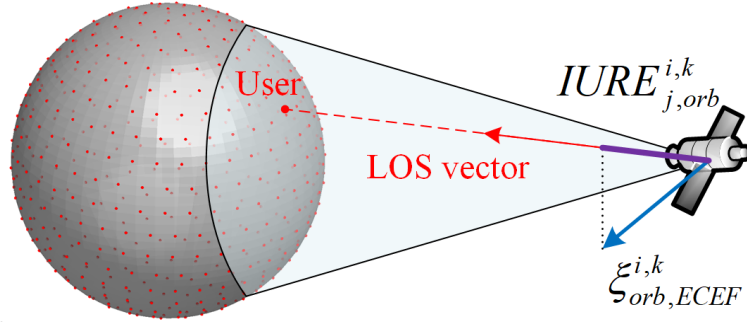


Figure 3.5: Demonstration of IURE over a mesh of 200 user locations.

each user positioned within the satellite's coverage footprint has a distinct LOS vector, the projected range error can vary with user locations. Therefore, the so-called instantaneous user range error (IURE) is defined. As shown in Figure 3.5, the orbit component of the IURE is calculated by projecting the orbit error $\xi_{orb,ECEF}^{i,k}$ to the LOS vector $\mathbf{x}_{i,j}^k$ from satellite i to receiver j

$$IURE_{j,orb}^{i,k} = \frac{\xi_{orb,ECEF}^{i,k} \mathbf{x}_{i,j}^k}{|\mathbf{x}_{i,j}^k|}, \quad (3.11)$$

where

$$\mathbf{x}_{i,j}^k = \mathbf{x}_j^k - \mathbf{p}^{i,k}. \quad (3.12)$$

The final IURE is calculated by adding the impacts of the clock error as follows:

$$IURE_j^{i,k} = IURE_{j,orb}^{i,k} - \tilde{\zeta}_{clk,ECEF}^{i,k}. \quad (3.13)$$

Note that the negative sign in Equation (3.13) accounts for the negative correlation between the clock error and the orbit error that results from the orbit and clock determination algorithm of GPS or Galileo [47].

The GPS Standard Positioning Service Performance Standard (GPS-SPS-PS) [60] defines the SISRE as the average contribution over all IURE values of users located inside the visibility cone of a satellite. However, the average value cannot reflect the worst-case scenario. Therefore, the globally averaged IURE is unsuitable for safety-critical applications.

Alternatively, the user projected error (UPE) is proposed by collecting the IURE of a large number of globally distributed users [61]. Each healthy satellite is visible to many but not all users at any given epoch; therefore, there will be multiple UPE values per satellite at each epoch (one for each user that has this satellite in view). Therefore, the UPE at epoch k can be formalized as follows:

$$UPE^{i,k} = \{IURE_j^{i,k} | \text{user } i \text{ is within the visibility cone of satellite } j \text{ at epoch } k\}. \quad (3.14)$$

Similar to the setting in [61], 200 evenly distributed user locations around the globe are selected to calculate the UPE, as shown in Figure 3.5. Finally, the SISRE of satellite i over a period is defined as

$$SISRE_{UPE}^i = \{UPE^{i,k} | \text{epoch } k \text{ is included in the given period}\}. \quad (3.15)$$

3.4 Nominal Performance Characterization

According to the GPS-SPS-PS [60], a GPS satellite is considered to have a major service failure when its average projected error exceeds $4.42 \times \sigma_{URA}$. A similar definition for Galileo satellites is given in the Galileo Open Service Service Definition Document (OS-SDD) [62], where a Galileo satellite is deemed faulty if the average projected error surpasses $4.17 \times \sigma_{URA}$. Table 3.1 displays the four identified GPS fault events during the study period (2020–2022), while Table 3.2 presents the two Galileo fault events identified within the same period. The data in these tables are sourced from [1], which offers a comprehensive analysis of satellite fault events. To characterize the nominal performance characterization of SIS errors, we exclude samples from the periods listed in Tables 3.1 and 3.2 in the following sections.

Table 3.1: GPS fault events between 2020 and 2022 (Taken from [1])

SVN	Date	UTC Time
53	February 7, 2021	17:00 – 21:15
69	March 2, 2021	07:10 – 08:15
73	September 24, 2021	11:55 – 14:15
58	October 2, 2022	15:10 – 15:55

Table 3.2: Galileo fault events between 2020 and 2022 (Taken from [1])

SVN	Date	UTC Time
210	September 5, 2021	05:43 – 6:02
210	April 29, 2022	01:00 – 01:12

3.4.1 Nominal performance of GPS SIS

Over a span of three years, a comprehensive analysis is conducted on a total of 30 GPS satellites, accounting for 2,999,386 orbit and clock nominal error samples. The folded cumulative distribution function (CDF) of GPS orbit and clock errors, compiled from all samples within this period, is depicted in Figure 3.6. The folded CDF depicts the second half of the CDF plot with $1-CDF$ for values of $x > m$, where m is the median of the distribution. As can be seen, most orbit and clock errors are located within ± 5 m. Among the three orbit error components, the along-track error shows the largest magnitude and dispersion, with the cross-track error following closely behind. However, both the along-track and cross-track errors have little contribution to the range error, as these two directions are almost perpendicular to the LOS vectors. The radial error, in contrast, closely aligns with the LOS vectors and has the largest contribution to the range error, showing the smallest error magnitude and dispersion among the orbit components. As another key contributor to range errors, the clock error shows a much larger error magnitude and dispersion than the radial error. Therefore, $SISRE_{UPE}$ is most similar to the clock error, as shown in Figure 3.6.

Figure 3.7 plots the folded CDF of $SISRE_{UPE}$ for each GPS satellite, where significant differences among satellites are observed. Some satellites, such as SVN 44, SVN 51, SVN 73, and SVN 65, exhibit large error magnitude and dispersion, with their maximum $SISRE_{UPE}$ exceeding 10 m. However, the $SISRE_{UPE}$ of most satellites is relatively small, which retains within the range of ± 5 m. Table 3.3 summarizes the standard deviation of the $SISRE_{UPE}$ for each satellite, which also suggests the difference among satellites. The mean of $SISRE_{UPE}$ for each satellite is also listed in Table 3.3, with the magnitude less than 5 cm for most satellites. Importantly, GPS precise orbit and clock accuracy stand between 3–4 cm. Since the $SISRE_{UPE}$ references the precise orbit and clock as the truth orbit and clock, its mean values are reasonably on the order of the precise orbit and clock accuracy. Over a long term period of several years, the $SISRE_{UPE}$ will eventually be nearly zero mean [25].

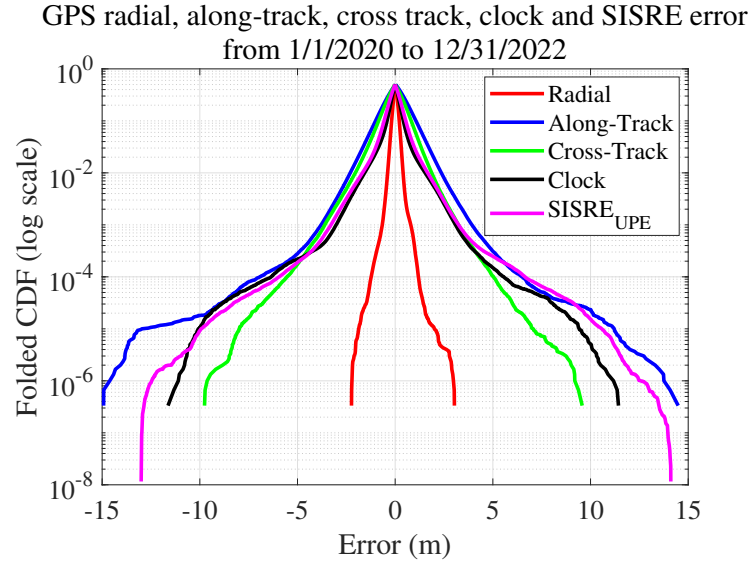


Figure 3.6: The folded CDF of GPS radial, along-track, cross-track, clock, and $SISRE_{UE}$ errors from January 1st, 2020 to December 31st, 2022

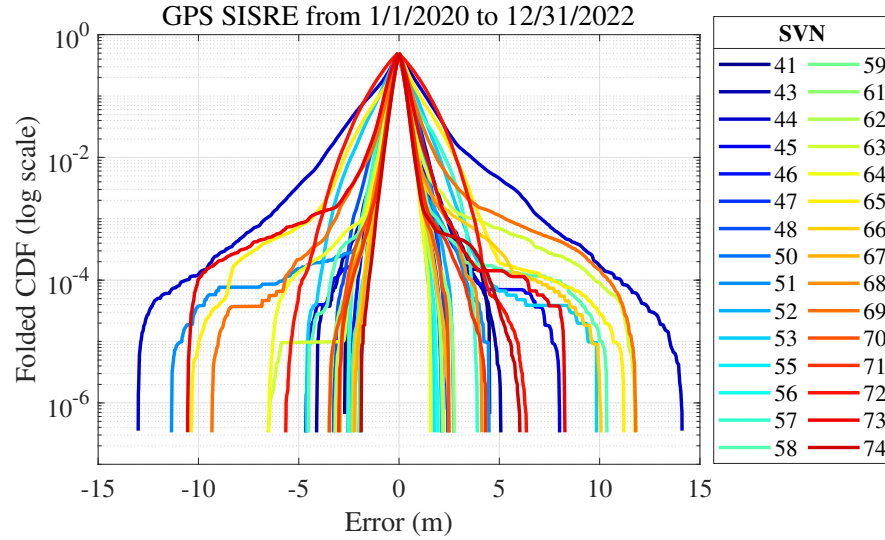


Figure 3.7: The folded CDF of GPS $SISRE_{UE}$ for individual satellites from January 1st, 2020 to December 31st, 2022.

Figure 3.8 shows the quantile-quantile (QQ) plot of the $SISRE_{UE}$ for each GPS satellite. The QQ plot shows the quantile of error distribution with the equivalent standard normal quantile, such that the Gaussian distributed error exhibits a

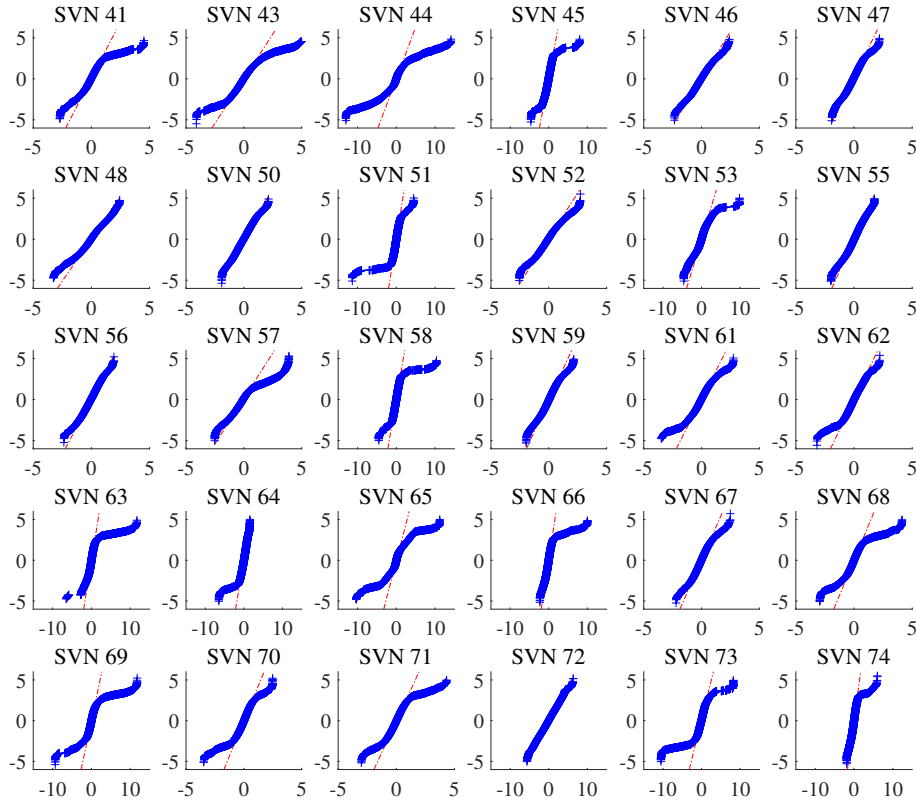


Figure 3.8: The QQ plot of GPS $SISRE_{UPE}$ for individual satellites from January 1st, 2020 to December 31st, 2022. The x-axis represents the quantile of error distribution (unit: meter) while the y-axis stands for the standard normal quantile (unit: meter).

straight line. Three categories of $SISRE_{UPE}$ distributions can be identified as follows: 1) Two-side heavy-tailed $SISRE_{UPE}$; 2) One-side heavy-tailed $SISRE_{UPE}$; and 3) Gaussian-liked $SISRE_{UPE}$. The category information is also provided in Table 3.3. For the two-side and one-side heavy-tailed $SISRE_{UPE}$, the Gaussian overbound used in the baseline ARAIM algorithm tends to provide conservative bounds, which potentially enlarges the protection level in the positioning domain. This issue is further discussed in Chapter 4, where a non-Gaussian overbounding method is developed to address this problem.

3.4.2 Nominal performance of Galileo SIS

For Galileo satellites, 6,836,848 orbit and clock error samples within 2020–2022 are used for analyzing their nominal SIS performance. Figure 3.9 shows the folded CDF of Galileo radial, along-track, cross-track, clock, and $SISRE_{UPE}$ errors, with all samples combined during the three-year period. The mean and standard deviation of these errors for individual satellites are summarized in Table 3.4. As shown in Figure 3.9, the radial and along-track errors have similar error magnitude and dispersion, while the cross-track error shows the smallest error magnitude and dispersion among the three orbit error components. This phenomenon is different from the GPS case. However, the relationship $\hat{\sigma}_{\text{rad}} < \hat{\sigma}_{\text{cross}} < \hat{\sigma}_{\text{along}}$ still holds in Galileo satellites, as shown in Table 3.4.

Notably, the folded CDF of Galileo orbit and clock errors exhibits an extremely narrow core, where most errors are located within ± 2 m. In the GPS case, most errors are located within ± 5 m. However, the dispersion of the orbit and clock errors of Galileo satellites is significantly larger than that of GPS satellites, with the maximum error magnitude reaching 25 m. These findings suggest that the nominal orbit and clock errors of Galileo satellites are usually smaller than those of GPS satellites, but Galileo satellites have larger worst-case nominal orbit and clock errors. Despite all these differences, we find a common point between Galileo and GPS satellites, i.e., the Galileo clock error also shows a larger error magnitude and dispersion than the orbit errors. Consequently, the $SISRE_{UPE}$ is similar to the clock error in the Galileo case.

The folded CDF of $SISRE_{UPE}$ for individual Galileo satellites is depicted in Figure 3.10 and the corresponding QQ plot is given in Figure 3.11. Two categories of $SISRE_{UPE}$ distributions can be identified as follows: 1) Two-side heavy-tailed $SISRE_{UPE}$ and 2) One-side heavy-tailed $SISRE_{UPE}$. Intuitively speaking, the tailedness of the Galileo $SISRE_{UPE}$ is much heavier than that of the GPS $SISRE_{UPE}$. However, the statistics of Galileo $SISRE_{UPE}$ in Table 3.4 suggests that the stan-

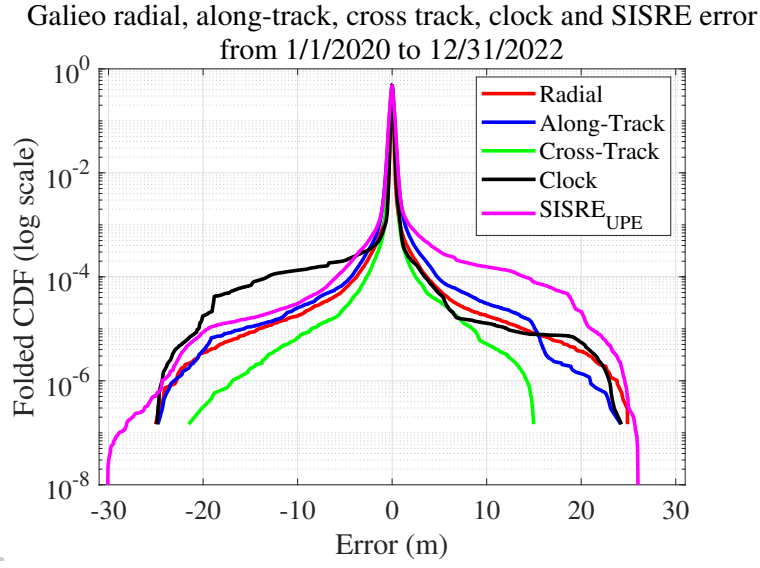


Figure 3.9: The folded CDF of Galileo radial, along-track, cross-track, clock, and $SISRE_{UPE}$ errors from January 1st, 2020 to December 31st, 2022.

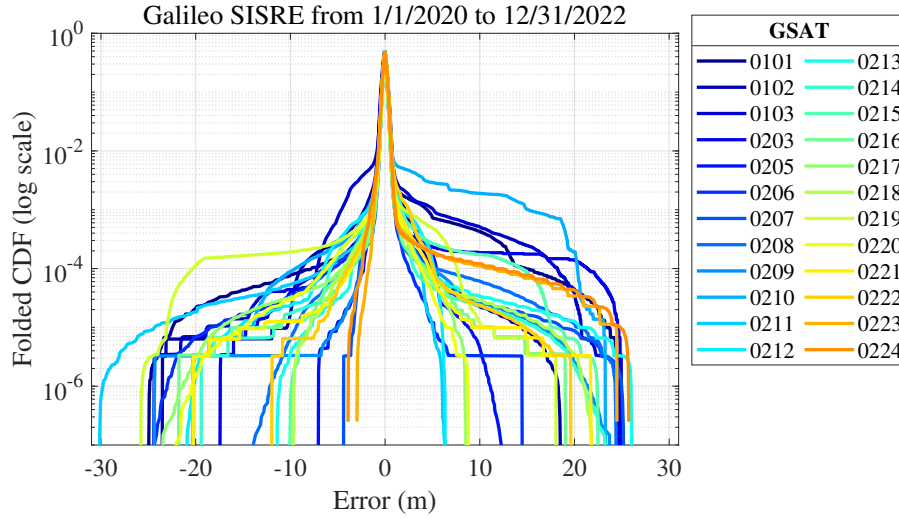


Figure 3.10: The folded CDF of Galileo $SISRE_{UPE}$ for individual satellites from January 1st, 2020 to December 31st, 2022.

dard deviation of the Galileo $SISRE_{UPE}$ is relatively smaller than that of the GPS $SISRE_{UPE}$. From Figure 3.9, we know that the nominal orbit and clock errors of Galileo satellites are usually smaller than those of GPS satellites, but Galileo satel-

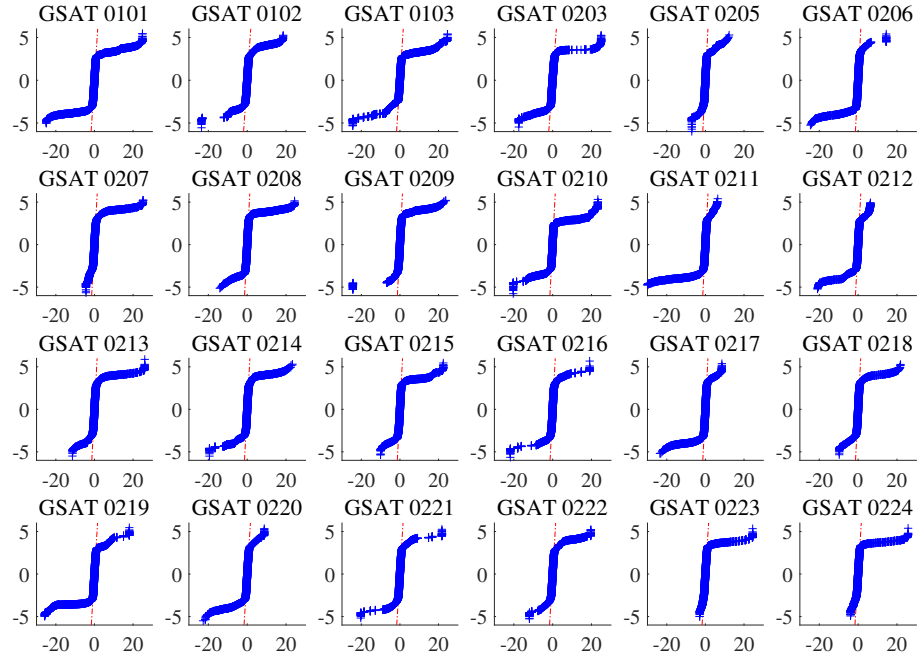


Figure 3.11: The QQ plot of Galileo $SISRE_{UPE}$ for individual satellites from January 1st, 2020 to December 31st, 2022. The x-axis represents the quantile of error distribution (unit: meter) while the y-axis stands for the standard normal quantile (unit: meter).

lites have larger worse-case nominal orbit and clock errors. Since the $SISRE_{UPE}$ is contributed by orbit and clock errors, the conclusion drawn from Figure 3.9 consequently extends to the $SISRE_{UPE}$. This explains why the Galileo $SISRE_{UPE}$ has a smaller standard deviation but exhibits heavier tails than the GPS $SISRE_{UPE}$. Finally, another important information in Table 3.4 is that the mean value of the Galileo $SISRE_{UPE}$ is nearly zero, which is similar to the GPS case.

Table 3.3: The mean and standard deviation of the radial, along-track, cross-track, clock, and $SISRE_{UPE}$ errors for individual GPS satellites from 1/1/2020 to

12/31/2022 (unit: cm)

SVN	Type ¹	Radial		Along-Track		Cross-Track		Clock		$SISRE_{UPE}$	
		mean	std	mean	std	mean	std	mean	std	mean	std
SVN41	O	0.31	11.25	-10.72	89.58	-5.51	54.09	0.82	33.31	-3.45	42.51
SVN43	T	0.29	11.15	10.69	86.72	-0.98	76.68	0.04	45.20	1.09	52.02
SVN44	T	3.20	23.42	6.99	106.60	0.31	92.18	0.53	123.26	-2.30	131.00
SVN45	O	3.39	15.96	-11.85	96.24	-4.50	86.58	-1.11	29.33	-4.67	42.82
SVN46	G	2.07	13.64	-5.09	113.00	3.37	55.05	0.22	34.55	-2.44	46.16
SVN47	G	3.88	13.48	-15.47	76.44	-1.03	60.87	-0.07	25.08	-6.14	36.59
SVN48	G	1.78	13.59	0.00	87.32	0.48	77.91	1.24	47.68	-2.97	55.04
SVN50	G	3.65	15.58	-6.56	78.43	2.76	61.92	-2.18	27.69	-1.99	40.48
SVN51	O	1.91	10.79	13.02	76.13	1.20	65.55	-1.15	30.26	1.33	38.70
SVN52	G	1.45	12.23	-0.71	72.77	-1.35	62.31	2.75	43.28	-4.44	49.91
SVN53	O	2.41	15.84	-2.24	112.50	-4.83	101.11	1.15	71.00	-4.49	80.98
SVN55	G	2.59	11.70	3.31	81.50	0.84	71.73	-2.95	21.74	1.08	34.04
SVN56	G	3.04	13.46	-18.12	83.12	0.29	63.62	-0.89	24.98	-4.66	37.62
SVN57	O	3.94	14.64	17.87	81.23	-3.01	61.31	-0.80	55.75	-0.96	62.42
SVN58	O	2.28	14.59	12.10	91.19	-1.93	82.14	-0.08	29.35	-0.70	40.14
SVN59	G	1.42	12.56	22.54	88.21	-2.79	67.24	-0.80	24.17	2.20	35.15
SVN61	T	2.64	14.07	7.15	108.17	-3.35	87.85	0.49	27.02	-2.55	40.05
SVN62	O	-2.20	18.26	10.02	96.69	-2.21	87.62	1.98	27.36	1.27	35.90
SVN63	T	0.73	18.00	4.39	84.49	0.13	68.36	-3.01	36.91	3.03	46.34
SVN64	O	0.35	17.93	0.06	84.66	0.05	59.14	-0.70	31.51	0.38	38.94
SVN65	T	-1.74	20.26	11.12	91.37	-4.58	75.78	-0.43	90.42	3.15	95.60
SVN66	O	-0.24	18.65	3.72	80.43	-1.58	68.31	1.68	28.22	-1.13	39.58
SVN67	G	2.42	17.38	-4.38	82.90	-0.54	70.36	-1.58	21.46	-1.51	33.14
SVN68	T	0.06	17.02	14.29	74.87	0.34	61.28	1.47	25.60	0.54	35.41
SVN69	T	-1.12	19.54	7.82	96.22	-0.90	90.78	-2.25	60.93	4.35	65.74
SVN70	T	-0.69	15.11	4.05	77.26	-3.43	65.33	-0.11	21.10	0.86	32.30
SVN71	T	-0.16	19.53	-2.51	80.22	-2.39	67.32	-2.20	22.21	1.65	36.05
SVN72	G	1.90	23.66	-15.36	104.69	-1.65	83.87	-0.20	115.60	-4.01	121.97
SVN73	T	-0.27	18.77	10.43	94.89	-1.81	81.16	8.46	55.07	-6.99	62.01
SVN74	O	-0.12	10.58	-5.73	83.41	0.29	49.19	-1.28	20.58	0.62	32.36

¹ “T”: Two-side heavy-tailed; “O”: One-side heavy-tailed; “G”: Gaussian-liked.

Table 3.4: The mean and standard deviation of the radial, along-track, cross-track, clock, and $SISRE_{UPE}$ errors for individual Galileo satellites from 1/1/2020 to 12/31/2022 (unit: cm)

SVN	Type ¹	Radial		Along-Track		Cross-Track		Clock		$SISRE_{UPE}$	
		mean	std	mean	std	mean	std	mean	std	mean	std
GSAT0101	T	-0.26	24.25	-1.16	29.02	-1.12	23.46	2.01	36.63	-2.11	48.13
GSAT0102	T	-0.54	20.66	-5.91	36.95	-1.18	25.97	2.13	21.35	-2.66	37.72
GSAT0103	T	-0.86	22.12	-2.77	36.93	-3.43	26.84	1.44	46.17	-1.50	57.17
GSAT0203	T	-0.42	19.08	-1.60	25.97	-1.00	20.55	4.70	37.61	-4.69	45.92
GSAT0205	O	-0.65	15.61	-1.03	26.81	-0.42	19.93	1.76	13.67	-1.34	27.12
GSAT0206	T	-0.36	17.27	-0.19	25.06	-0.90	21.67	1.56	14.90	-1.38	27.76
GSAT0207	O	-0.79	17.55	-3.02	27.81	-1.01	20.34	1.66	15.29	-1.48	29.75
GSAT0208	T	-0.82	18.66	-3.28	27.98	-1.15	20.41	1.27	13.66	-1.12	29.30
GSAT0209	T	-0.73	16.81	-0.71	26.93	-1.99	19.96	1.13	13.81	-0.81	27.30
GSAT0210	T	-0.21	19.10	-2.73	37.15	-1.44	24.87	-0.05	78.24	-0.36	82.48
GSAT0211	O	-0.57	18.69	-0.86	28.42	-1.84	21.14	1.46	14.74	-1.32	30.73
GSAT0212	O	-0.81	18.20	-2.41	27.92	-1.27	20.75	1.22	15.40	-0.96	32.05
GSAT0213	T	-0.91	18.62	-2.36	29.26	-1.65	20.72	0.60	14.15	-0.30	29.95
GSAT0214	T	-0.80	18.34	-0.10	26.25	-1.97	20.70	1.03	15.38	-0.55	29.92
GSAT0215	T	-0.71	17.87	-0.64	27.90	-0.78	19.86	0.80	22.90	-0.31	33.80
GSAT0216	T	-0.55	15.46	-1.17	27.30	-1.14	19.06	1.15	15.50	-0.96	27.89
GSAT0217	T	-0.54	18.25	-1.60	26.42	-0.31	21.54	1.47	13.08	-1.23	27.23
GSAT0218	T	-0.80	18.44	-0.37	26.21	-1.25	19.89	1.41	12.82	-0.87	27.81
GSAT0219	T	-0.36	23.20	-0.18	43.44	-0.94	25.00	1.49	29.24	-1.30	43.80
GSAT0220	T	-0.45	20.85	-2.04	29.33	-0.56	20.64	-1.83	14.93	1.87	31.99
GSAT0221	T	-0.18	18.35	-2.78	28.51	-1.64	20.79	1.60	17.06	-2.09	31.33
GSAT0222	T	-0.29	19.07	-2.71	28.23	-0.08	21.99	1.48	15.74	-1.63	31.92
GSAT0223	O	-0.85	24.27	-0.26	26.59	-1.64	20.28	1.65	15.45	-1.10	35.24
GSAT0224	O	-0.83	25.09	-0.75	30.81	-1.63	20.99	1.25	15.98	-0.80	35.56

¹ “T”: Two-side heavy-tailed; “O”: One-side heavy-tailed.

Chapter 4

Principal Gaussian Overbound for Heavy-tailed Error Distribution

To meet the stringent navigation requirements, precise modeling of the measurement error distribution using experimental data is essential. Chapter 3 accurately characterizes the SISRE distribution of GPS and Galileo satellites, which use a huge amount of data. However, it is impossible to transmit all these data to users in satellite or augmentation systems due to limited-bandwidth communication. To address this issue, a conservative representation of the error distribution known as the overbound should be employed. For heavy-tailed errors, such as the SISRE, additional challenges are imposed to overbounds. A sharp yet conservative overbound is preferred to support integrity monitoring applications due to availability and continuity constraints.

This chapter proposes the Principal Gaussian overbound (PGO) for heavy-tailed error distributions by leveraging the characteristics of the GMM. Section 4.1 gives a brief review of the existing overbounding methods. In Section 4.2, the GMM and its application on error fitting and overbounding research are illustrated, which provides the theoretical foundation for the development of PGO in Section 4.3. Section 4.4 compares the bounding performance of the PGO with the two-step Gaussian overbound

and the Gaussian-Pareto overbound on GPS and Galileo SISRE. Finally, Section 4.5 constructs the nominal error bounds of code IF combination based on the proposed PGO.

4.1 Review of Overbounding Methods

The first true overbound, known as the CDF overbound, was introduced by DeCleene in 2000 [63]. It is defined as having more tail mass than the error distribution. To facilitate the range-to-position projection and simplify the communication and computation of error bounds, the zero-mean Gaussian model is adopted as the underlying form of the CDF overbound. Since then, the Gaussian CDF overbound and its variants have dominated the overbounding research and are taken as the basis for integrity analysis. Nevertheless, DeCleene's method requires certain shape constraints on the overbounding distribution, including symmetric and unimodality.

To relax the constraints in the DeCleene's method, Rife et al. proposed the paired CDF overbound that uses two Gaussian distributions with non-zero mean to overbound the left and right regions separately [64]. However, its stringent requirement on bounding both regions inflates the standard deviation or enlarges the biases in the Gaussian overbound [65]. This problem is relieved by allowing the total mass of the overbound distribution to be greater than one, which is known as the excess-mass CDF overbound [66]. More recently, Blanch et al. proposed the two-step Gaussian overbounding method, significantly reducing the bias in the overbound distribution [65].

Nonetheless, these Gaussian-based overbounding methods fail to truly overbound the heavy-tailed distributions, a frequent occurrence in pseudorange errors influenced by multipath [67]. This is because heavy-tailed distributions have tails that are not exponentially bounded, which makes them impossible to be overbounded by a Gaussian distribution that only possesses an exponential tail [22, 68]. It is natural

to apply different models for the core and tail regions of a distribution. Rife et al. proposed the Gaussian core overbound, which utilizes a Gaussian distribution for the core bounding and an implicit distribution for the tail bounding [22]. More recently, Larson et al. proposed the Gaussian-Pareto overbound to tightly bound the tails by utilizing the extreme value theory [68]. However, the overbounding property of the Gaussian-Pareto overbound through convolution remains unclear, which currently limits its applications.

In the remainder of this section, the construction process of the two-step Gaussian overbound and the Gaussian-Pareto overbound is illustrated, as both the two methods will be used for comparison with the proposed method in the following sections.

4.1.1 Two-step Gaussian overbound

The two-step Gaussian overbound [65], as the name suggests, involves two steps of construction. In the first step, a piecewise uniform, symmetric, and unimodal distribution $G_{\text{su}}(x)$ is determined through an ad hoc approach:

$$G_v(x) \leq G_{\text{su}}(x) \quad \forall x. \quad (4.1)$$

In the second step, the left-hand side overbound is determined by finding the minimum σ_L that satisfies

$$\int_{-\infty}^x f_N(x; -b_L, \sigma_L) dx \geq G_{\text{su}}(x) \quad \forall x \leq 0. \quad (4.2)$$

Equation (4.2) relaxes the constraint in paired overbound [64] by only CDF bounding the left-hand side of the intermediate distribution. The right-hand side overbound is obtained by repeating the above process on the mirror image of the error distribution, i.e., $G_v(-x)$. The right-hand side overbound has the CDF form as follows:

$$\int_{-\infty}^x f_N(x; b_R, \sigma_R) dx \quad \forall x > 0. \quad (4.3)$$

Although the two-step Gaussian overbounding method can significantly reduce the bias compared to the conventional paired overbounding methods [65], its overbound

for the heavy-tailed distribution can still be conservative, which will be shown through numerical experiments in Section 4.4.

4.1.2 Gaussian-Pareto overbound

The Gaussian-Pareto overbound is a realization of the core overbounding concept [22]. Specifically, the core overbounding concept decomposes the error distribution into two fractions, including the core part $G_{v,\text{core}}(x)$ and the tail part $G_{v,\text{tail}}(x)$ as follows,

$$G_v(x) = G_{v,\text{core}}(x) + G_{v,\text{tail}}(x), \quad (4.4)$$

where $G_{v,\text{core}}(x)$ and $G_{v,\text{tail}}(x)$ are defined in terms of the error distribution PDF $f_v(x)$ as follows:

$$G_{v,\text{core}}(x) = \begin{cases} \int_{-\infty}^x f_v(x) dx & |x| \leq T \\ 0 & |x| > T \end{cases} \quad (4.5a)$$

$$G_{v,\text{tail}}(x) = \begin{cases} 0 & |x| \leq T \\ \int_{-\infty}^x f_v(x) dx & |x| > T \end{cases}, \quad (4.5b)$$

where T is the core-tail transition point. The core overbound of $G_v(x)$ is defined as

$$G_{ov}(x) = \hat{G}_{ov,\text{ex}}(x) + \hat{G}_{ov,\text{im}}(x), \quad (4.6)$$

where $\hat{G}_{ov,\text{ex}}(x)$ and $\hat{G}_{ov,\text{im}}(x)$ are CDFs scaled by a factor within the range of zero to one. In general, $\hat{G}_{ov,\text{ex}}(x)$ is an explicit function that bounds the worse-case CDF of the core part of the error distribution,

$$G_{v,\text{core}}(x) \leq \hat{G}_{ov,\text{ex}}(x) \quad \forall -T \leq x \leq 0 \quad (4.7a)$$

$$G_{v,\text{core}}(x) \geq \hat{G}_{ov,\text{ex}}(x) \quad \forall 0 < x \leq T, \quad (4.7b)$$

and $G_{ov,im}(x)$ is an implicit function (i.e., no assumptions are made on its shape) that bounds the tail part of the error distribution,

$$G_{v,tail}(x) \leq \hat{G}_{ov,im}(x) \quad \forall x < -T \quad (4.8a)$$

$$G_{v,tail}(x) \geq \hat{G}_{ov,im}(x) \quad \forall x > T. \quad (4.8b)$$

As a realization of the core overbounding concept, the Gaussian-Pareto overbound [68] adopts the generalized Pareto distribution for the tail bounding. The tail of the Gaussian-Pareto overbound can be formalized as

$$\hat{G}_{GP,tail}(x) = \begin{cases} G_v(u_L) - G_{GPD}(u_L - x) G_v(u_L) & \forall x < u_L \\ G_{GPD}(x - u_R) (1 - G_v(u_R)) + G_v(u_R) & \forall x > u_R \end{cases}, \quad (4.9)$$

where $u_L < 0$ is the left core-tail transition point, $u_R > 0$ is the right core-tail transition point, and $G_{GPD}(x; \mu, \beta, \gamma)$ is the CDF of the generalized Pareto distribution with a location parameter, u , a scale parameter, β , and a shape parameter, γ [69], as shown below:

$$G_{GPD}(x; u, \beta, \gamma) = \begin{cases} 1 - \left(1 + \frac{\gamma(x-u)}{\beta}\right)^{-\frac{1}{\gamma}} & \gamma \neq 0 \\ 1 - \exp\left(-\frac{x-u}{\beta}\right) & \gamma = 0 \end{cases}. \quad (4.10)$$

The Gaussian-Pareto overbound stems from the extreme value theory, which shows that almost all normalized continuous probability distributions asymptotically approach a generalized Pareto distribution [70]. This property enables the generalized Pareto distribution to provide a true overbound for data far beyond the end of the empirical distributions, as required in DeCleene's theorem [63]. Although the Gaussian-Pareto overbounding method offers the potential to tightly overbound the tail distribution, it is unclear whether the Gaussian-Pareto overbound can maintain the overbounding property through convolution, which is crucial for deriving pseudorange-level requirements from the position-domain integrity requirement [63].

4.2 Gaussian Mixture Model

GMM is a statistical technique that plays a crucial role in error modeling, particularly in capturing heavy-tailed distributions that are commonly observed in real-world scenarios [61, 71, 72]. A GMM represents a probability distribution as a weighted combination of multiple Gaussian distributions, each representing a component of the mixture. In particular, the bimodal Gaussian mixture model (BGMM) is of significant research interest as it effectively balances simplicity and flexibility, enabling precise modeling of both the core and tails of heavy-tailed distributions while avoiding overfitting and excessive parameters [71, 73]. The PDF of a zero-mean BGMM for heavy-tailed distribution modeling can be formalized as

$$f(x) = p_1 f_N(x; 0, \sigma_1) + (1 - p_1) f_N(x; 0, \sigma_2) , \quad (4.11)$$

where $f_N(x; 0, \sigma_1)$ and $f_N(x; 0, \sigma_2)$ are the PDF of the first and the second Gaussian component, σ_1 and σ_2 are the corresponding standard deviations, and p_1 and $1 - p_1$ are the mixing weight of the two Gaussian components, respectively. In this work, it is assumed that $\sigma_1 < \sigma_2$ and $p_1 \in (0.5, 1)$, indicating that the Gaussian component with the smaller standard deviation is selected as the 1st Gaussian component and exhibits a higher mixing weight. The parameters in a GMM are usually estimated by employing maximum likelihood estimation (MLE). The expectation-maximization (EM) algorithm [74] is usually adopted to obtain the MLE. A concise summary of the EM algorithm is provided in Appendix A.3.

Due to the appealing nature of GMM, research on GMM-based overbounds emerged as early as 2001 and has received increased attention in recent years [67]. Shively compared GMM overbounding methods with Gaussian and exponential overbounding methods in terms of the resulting inflation factor, suggesting that the GMM method may align most closely with the manner in which errors are present in actual data [67]. Lee used the GMM as an empirical distribution to model the ground facility error distribution, showing the potential of GMM to reduce the inflation factor and the

PL [73]. However, the PL is calculated by a Gaussian overbound of the GMM, which inevitably increases the conservatism. To tackle this issue, Blanch et al. constructed a BGMM for pseudorange error modeling in the worst case [71]. They proved that the posterior position density is also GMM and derived the PL by integrating the posterior position density. Nevertheless, the computation of the PL involves a large number of matrix inversions, which cause a considerable computation burden.

These studies usually obtain the GMM from the worst-case analysis [71] or sample data fitting [74, 75]. However, a rigorous method to determine the least conservative GMM overbound of the error distribution has not been developed. The crux of this issue lies in the difficulty of establishing overbounding relationships between two GMMs. In Gaussian scenarios, a Gaussian with a larger standard deviation can always bound the Gaussian with a smaller standard deviation. However, in the context of GMMs, the increasing parameters can provide more flexibility in shaping the distribution, which in turn makes it more difficult to assess whether one GMM's distribution covers the entire range of the other GMM's distribution. Therefore, this work explores an alternative way to utilize GMM to bound the heavy-tailed error distribution. We propose the PGO, which is proven to be the true CDF overbound of the GMM distribution.

4.3 Principal Gaussian Overbound for the Heavy-Tailed Error Distribution

4.3.1 Membership weight analysis

The process of generating samples from a GMM can be seen as simultaneously generating samples from multiple Gaussian distributions according to their respective weights. For example, given a K -component GMM with mixture weights of p_1 ,

$p_2, \dots, p_K$, the sample generated from it at time t is drawn from the 1st component with probability p_1 , drawn from the 2nd component with probability p_2 , and etc. From the perspective of statistical inference, we can also estimate the likelihood of a sample belonging to a specific Gaussian component, which is known as the membership weight. The membership weight indicates the posterior probability of a data point being generated from each component, which can be derived by utilizing the Bayes' Theorem [76]. For a given observation x , we define the allocation variable $c = \{1, 2, \dots, K\}$ that marks the Gaussian component from which x is generated. Then, the mixture weight of the k th Gaussian component can be interpreted as the prior probability of the allocation variable that equals k , as shown below:

$$p_k = P(c = k) . \quad (4.12)$$

For a given $c = k$, the probability of generating x (i.e., the likelihood) is

$$P(x|c = k) = f_N(x; b_k, \sigma_k) . \quad (4.13)$$

According to Bayes' Theorem, the posterior probability is given by

$$s_k = P(c = k|x) = \frac{p_k f_N(x; b_k, \sigma_k)}{\sum_{k=1}^K p_k f_N(x; b_k, \sigma_k)} , \quad (4.14)$$

where s_k is the membership weight of the k th Gaussian component for the observation x .

In this work, we mainly focus on the zero-mean BGMM in Equation (4.11) due to its good balance of simplicity and flexibility. The membership weights s_1 and s_2 of the zero-mean BGMM can be written as

$$s_1(x) = \frac{p_1 f_N(x; 0, \sigma_1)}{p_1 f_N(x; 0, \sigma_1) + (1 - p_1) f_N(x; 0, \sigma_2)} \quad (4.15a)$$

$$\begin{aligned} s_2(x) &= \frac{(1 - p_1) f_N(x; 0, \sigma_2)}{p_1 f_N(x; 0, \sigma_1) + (1 - p_1) f_N(x; 0, \sigma_2)} \\ &= 1 - s_1(x) . \end{aligned} \quad (4.15b)$$

By substituting the probability density function (PDF) equation of a Gaussian distribution into Equation (4.15a),

$$\begin{aligned}
 s_1(x) &= \frac{\frac{p_1}{\sigma_1} \exp\left(-\frac{1}{2} \left(\frac{x}{\sigma_1}\right)^2\right)}{\frac{p_1}{\sigma_1} \exp\left(-\frac{1}{2} \left(\frac{x}{\sigma_1}\right)^2\right) + \frac{1-p_1}{\sigma_2} \exp\left(-\frac{1}{2} \left(\frac{x}{\sigma_2}\right)^2\right)} \\
 &= \frac{1}{1 + \frac{1-p_1}{p_1} \frac{\sigma_1}{\sigma_2} \exp\left(\frac{\sigma_2^2 - \sigma_1^2}{2\sigma_1^2 \sigma_2^2} x^2\right)}.
 \end{aligned} \tag{4.16}$$

Since $\sigma_2 > \sigma_1$, $\exp\left(\frac{\sigma_2^2 - \sigma_1^2}{2\sigma_1^2 \sigma_2^2} x^2\right)$ will be a symmetric convex function in terms of x ; Hence, $s_1(x)$ will be a symmetric concave function. Since $s_2(x) = 1 - s_1(x)$, $s_2(x)$ will consequently be a symmetric convex function. Figure 4.1 plots $s_1(x)$ and $s_2(x)$ in two settings of GMMs, including $p_1 = 0.9, \sigma_1 = 0.5, \sigma_2 = 0.7$ in Figure 4.1a and $p_1 = 0.9, \sigma_1 = 0.5, \sigma_2 = 1.5$ in Figure 4.1c. In both cases, $s_1(x)$ has large values when x is located at the central region of the BGMM, and the value decreases dramatically when x goes far away from the center. However, $s_2(x)$ shows an opposite trend that the largest value of $s_2(x)$ is located at the tail region. These trends indicate the dominance of each Gaussian component in different regions of the BGMM, which provides the theoretical foundation for the dominance partition illustrated in the next section.

4.3.2 Dominance partition

It is straightforward to use the intersection points of $s_1(x)$ and $s_2(x)$ to partition the core and tail regions of the zero-mean BGMM. The intersection points $x_{\text{intersect}}^L$ and $x_{\text{intersect}}^R$ can be found by solving $s_1(x) = s_2(x)$ as

$$x_{\text{intersect}}^L = -\sqrt{\frac{2\sigma_1^2 \sigma_2^2}{\sigma_2^2 - \sigma_1^2} \ln \frac{p_1 \sigma_2}{(1-p_1) \sigma_1}} \tag{4.17a}$$

$$x_{\text{intersect}}^R = \sqrt{\frac{2\sigma_1^2 \sigma_2^2}{\sigma_2^2 - \sigma_1^2} \ln \frac{p_1 \sigma_2}{(1-p_1) \sigma_1}}. \tag{4.17b}$$

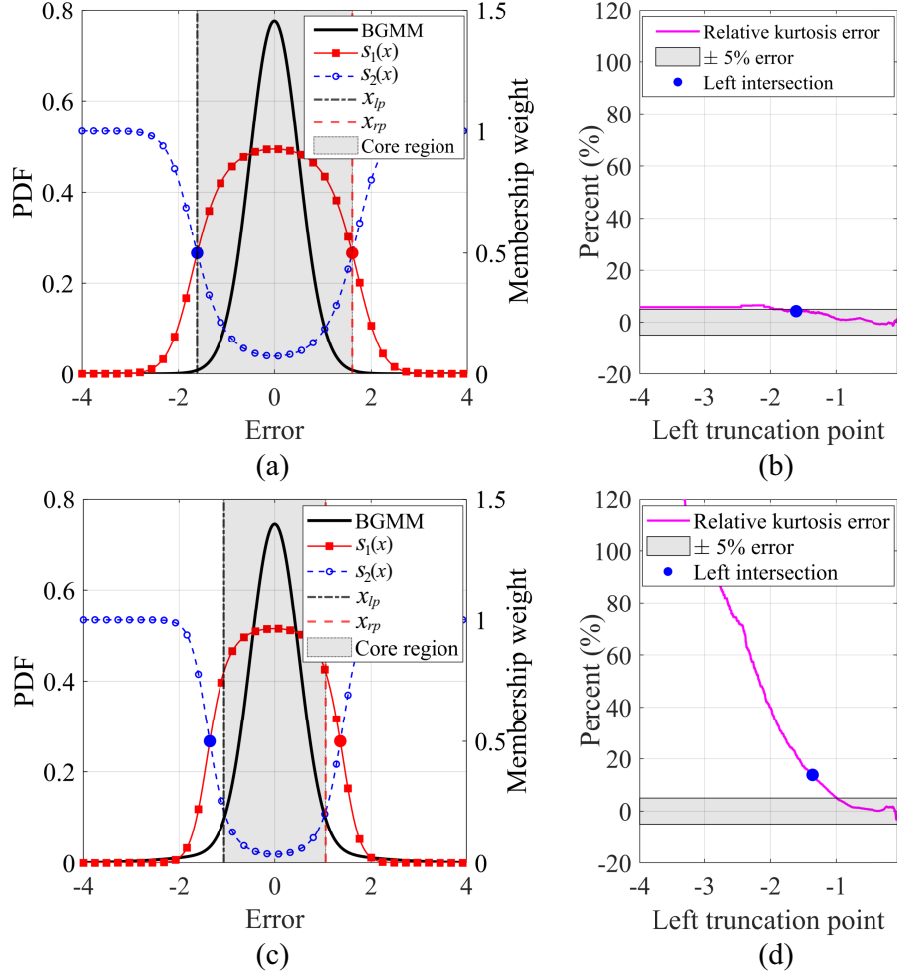


Figure 4.1: (a) Membership weights and (b) the relative kurtosis error of a zero-mean BGMM with $p_1 = 0.9$, $\sigma_1 = 0.5$, and $\sigma_2 = 0.7$; (c) Membership weights and (d) the relative kurtosis error of a zero-mean BGMM with $p_1 = 0.9$, $\sigma_1 = 0.5$, and $\sigma_2 = 1.5$. x_{lp} and x_{rp} are the core-tail transition points. The two intersection points of $s_1(x)$ and $s_2(x)$ are marked with blue and red solid points.

According to the convex property of $s_2(x)$, the value of $s_2(x)$ will exceed 0.5 and increase monotonically as the data point moves away from $x_{\text{intersect}}^R$ to ∞ (or from $x_{\text{intersect}}^L$ to $-\infty$). Meanwhile, $s_1(x)$ will decrease monotonically and eventually go to zero. This trend indicates that the 2nd Gaussian component dominates the region $x \in (-\infty, x_{\text{intersect}}^L] \cup [x_{\text{intersect}}^R, \infty)$.

However, the dominance relationship in the region $x \in [x_{\text{intersect}}^L, x_{\text{intersect}}^R]$ is more complicated than that in the region $x \in (-\infty, x_{\text{intersect}}^L] \cup [x_{\text{intersect}}^R, \infty)$. Although $s_2(x)$ monotonically decreases when x gradually moves to the center, $s_2(x)$ will not be reduced to zero, which indicates that the impacts of the 2nd Gaussian component in the region $x \in [x_{\text{intersect}}^L, x_{\text{intersect}}^R]$ cannot be ignored. To quantify the impacts of the 2nd Gaussian component on the tailedness of the BGMM distribution, we calculate the kurtosis, a measure of tailedness, of the doubly truncated zero-mean BGMM and compare it with the kurtosis of doubly truncated standard normal distribution [77,78]. Specifically, we randomly generate $N_t = 10,000$ samples from the zero-mean BGMM and truncate samples smaller than x_t^L or larger than x_t^R , where $x_t^L < 0$ and $x_t^R > 0$ are the truncation points. The truncation rate γ_t is defined by

$$\gamma_t = 1 - \frac{n_t}{N_t}, \quad (4.18)$$

where n_t is the number of samples in $[x_t^L, x_t^R]$. Then we calculate the kurtosis of samples within $[x_t^L, x_t^R]$ by

$$k_{\text{BGMM}}(x_t^L) = \frac{\frac{1}{n_t} \sum_{x_i \in [x_t^L, x_t^R]} (x_i - \bar{x})^4}{\left[\frac{1}{n_t} \sum_{x_i \in [x_t^L, x_t^R]} (x_i - \bar{x})^2 \right]^2}, \quad (4.19)$$

where $\bar{x}$ is the mean of samples within $[x_t^L, x_t^R]$. The corresponding truncation points of the standard normal distribution can be calculated by

$$\begin{aligned} x_{\text{normal}}^L &= Q^{-1}\left(\frac{\gamma_t}{2}\right) \\ x_{\text{normal}}^R &= -x_{\text{normal}}^L, \end{aligned} \quad (4.20)$$

where $Q^{-1}(\cdot)$ is the quantile function of the standard normal distribution. Equation (4.20) ensures that the doubly truncated standard normal distribution has the same truncation rate as the doubly truncated zero-mean BGMM. Similarly, we generate $N_t = 10,000$ samples from the standard normal distribution and calculate the kurtosis of the truncated distribution by setting x_{normal}^L and x_{normal}^R as the truncation points. The kurtosis of the doubly truncated standard normal distribution is denoted as

$k_{\text{normal}}(x_t^L)$. We then calculate the relative error between $k_{\text{BGMM}}(x_t^L)$ and $k_{\text{normal}}(x_t^L)$ by

$$e_k(x_t^L) = \frac{k_{\text{BGMM}}(x_t^L) - k_{\text{normal}}(x_t^L)}{k_{\text{normal}}(x_t^L)}, \quad (4.21)$$

and plot its value against x_t^L in Figure 4.1.

Figure 4.1b shows the relative kurtosis error in the case of BGMM with $p_1 = 0.9$, $\sigma_1 = 0.5$, and $\sigma_2 = 0.7$. The relative kurtosis error is within the $\pm 5\%$ error region when $x_t^L \geq x_{\text{intersect}}^L$, indicating that samples in the core region $x \in [x_{\text{intersect}}^L, x_{\text{intersect}}^R]$ show similar tailedness with the truncated standard normal distribution. This indicates that $x_{\text{intersect}}^L$ could be a good core-tail transition point since the core region $x \in [x_{\text{intersect}}^L, x_{\text{intersect}}^R]$ is less affected by the 2nd Gaussian component. However, the situation becomes different in the case of BGMM with $p_1 = 0.9$, $\sigma_1 = 0.5$, and $\sigma_2 = 1.5$, as shown in Figure 4.1d. The relative kurtosis error is 14 % when $x_t^L = x_{\text{intersect}}^L = -1.36$, and this value slowly decreases to 5 % until x_t^L increases to -1 . If we adopt $x_{\text{intersect}}^L$ as the core-tail transition point, the core region could be severely affected by the 2nd Gaussian component, making it difficult to distinguish the dominance relationship in the core region. Therefore, it is more beneficial to use the truncation point with 5 % relative kurtosis error as the core-tail transition point. In summary, we use the following rules to choose the core-tail transition point

$$x_{lp} = \begin{cases} x_{\text{intersect}}^L & \text{if } |e_k(x_{\text{intersect}}^L)| \leq \alpha \\ x_t^L \text{ s.t. } e_k(x_t^L) = \alpha & \text{if } |e_k(x_{\text{intersect}}^L)| > \alpha \end{cases}, \quad (4.22)$$

$$x_{rp} = -x_{lp}, \quad (4.23)$$

where α is the partition parameter. We define $[x_{lp}, x_{rp}]$ as the core region and $(-\infty, x_{lp}] \cup [x_{rp}, \infty)$ as the tail region. In this work, we choose $\alpha = 0.05$.

4.3.3 Principal Gaussian overbound

In this section, we propose the PGO based on the scaling and shifting of the dominant Gaussian component of the BGMM in the tail and core regions. The PDF of the PGO takes the following form:

$$f_o(x) = \begin{cases} (1+k)(1-p_1)f_N(x;0,\sigma_2) & |x| > x_{rp} \\ p_1f_N(x;0,\sigma_1) + c & |x| \leq x_{rp} \end{cases}, \quad (4.24)$$

where k is the scaling parameter, and c is the shifting parameter. The remainder of this section illustrates the construction of PGO.

4.3.3.1 Tail region bounding

In the tail region, the BGMM is dominated by the 2nd Gaussian component. Therefore, the 2nd Gaussian component, along with its mixture weight, is taken as the basis for bounding the tail region. In addition, compensation is needed to account for the contribution of the 1st component in the tail region. An intuitive approach of the CDF overbound at the left tail region is shown as follows:

$$G_o^L(x) = p_1G(x_{lp};0,\sigma_1) + (1-p_1)G(x;0,\sigma_2) \quad \forall x < x_{lp}, \quad (4.25)$$

where $G(x;0,\sigma_1)$ and $G(x;0,\sigma_2)$ are the CDF of the 1st and 2nd Gaussian component, respectively; and $G(x_{lp};0,\sigma_1)$ is the CDF of the 1st Gaussian component evaluated at x_{lp} . Let $G(x)$ be the CDF of the BGMM defined in Equation (4.11),

$$\begin{aligned} G_o^L(x) - G(x) &= G_o^L(x) - (p_1G(x;0,\sigma_1) + (1-p_1)G(x;0,\sigma_2)) \\ &= p_1(G(x_{lp};0,\sigma_1) - G(x;0,\sigma_1)). \end{aligned} \quad (4.26)$$

Since $x < x_{lp}$, we have $G(x_{lp};0,\sigma_1) - G(x;0,\sigma_1) > 0$. Therefore, $G_o^L(x)$ is the CDF overbound of the BGMM at the left tail region. However, the CDF overbound in Equation (4.25) includes a constant term, $p_1G(x_{lp};0,\sigma_1)$, which is defined in an

unbounded interval. This poses a challenge in deriving the PDF of the overbound distribution, which is necessary for convolution purposes.

We notice that $p_1 G(x_{lp}; 0, \sigma_1)$ in Equation (4.25) is a constant term, which could be compensated by inflating the weight of the 1st Gaussian component, i.e., $G(x; 0, \sigma_2)$ in Equation (4.25). Therefore, we introduce the scaling parameter k into Equation (4.25) as follows:

$$G_o^L(x) = (1 + k)(1 - p_1) G(x; 0, \sigma_2) \quad \forall x < x_{lp}. \quad (4.27)$$

We need to determine the value of k so that Equation (4.27) is a CDF overbound. Let Equations (4.25) and (4.27) produce the same value at x_{lp} , the value of k can be determined by

$$k = \frac{p_1 G(x_{lp}; 0, \sigma_1)}{(1 - p_1) G(x_{lp}; 0, \sigma_2)}. \quad (4.28)$$

Appendix A.4 gives proof that $G_o^L(x)$ in Equation (4.27) is the CDF overbound at the left tail region. The PDF of the overbound distribution at the left tail region can be derived by taking the derivative of Equation (4.27) as

$$f_o^L(x) = (1 + k)(1 - p_1) f_N(x; 0, \sigma_2) \quad \forall x < x_{lp}. \quad (4.29)$$

Similarly, the PDF of the overbound distribution at the right tail region can be written as

$$f_o^R(x) = (1 + k)(1 - p_1) f_N(x; 0, \sigma_2) \quad \forall x > x_{rp}. \quad (4.30)$$

4.3.3.2 Core region bounding

In the core region, the BGMM is dominated by the 1st Gaussian component, as illustrated in Section 4.3.2. Nevertheless, the contribution of the 2nd Gaussian component to the probability distribution in the core region is not negligible. Therefore, we introduce a constant term c to compensate for such a contribution when developing the overbound distribution in the core region, as shown below:

$$f_o^{\text{core}}(x) = p_1 f_N(x; 0, \sigma_1) + c \quad \forall x_{lp} \leq x \leq x_{rp}, \quad (4.31)$$

To determine the value of c , we calculate the CDF of $f_o^{\text{core}}(x)$ through integration as follows:

$$G_o^{\text{core}}(x) = \int_{-\infty}^{x_{lp}} f_o^{\text{L}}(x) dx + \int_{x_{lp}}^x f_o^{\text{core}}(x) dx \quad (4.32a)$$

$$= G_o^{\text{L}}(x_{lp}) + c(x - x_{lp}) + p_1(G(x; 0, \sigma_1) - G(x_{lp}; 0, \sigma_1)) \quad (4.32b)$$

$$= p_1 G(x; 0, \sigma_1) + c(x - x_{lp}) + (1 - p_1) G(x_{lp}; 0, \sigma_2) . \quad (4.32c)$$

Let $G_o^{\text{core}}(0) = G(0)$, the value of c is determined by

$$c = \frac{(1 - p_1)(G(x_{lp}; 0, \sigma_2) - 0.5)}{x_{lp}} . \quad (4.33)$$

To prove that $G_o^{\text{core}}(x)$ is the CDF overbound in the core region, we calculate the difference between $G_o^{\text{core}}(x)$ and $G(x)$ as follows:

$$\begin{aligned} \Delta G(x) &= G_o^{\text{core}}(x) - G(x) \\ &= \left[p_1 G(x; 0, \sigma_1) + c(x - x_{lp}) + (1 - p_1) G(x_{lp}; 0, \sigma_2) \right] \\ &\quad - \left[p_1 G(x; 0, \sigma_1) + (1 - p_1) G(x; 0, \sigma_2) \right] \\ &= c(x - x_{lp}) - (1 - p_1) G(x; 0, \sigma_2) + (1 - p_1) G(x_{lp}; 0, \sigma_2) . \end{aligned} \quad (4.34)$$

The first and second derivatives of $\Delta G(x)$ can be obtained by

$$\Delta G(x)' = c - (1 - p_1) f_N(x; 0, \sigma_2) \quad (4.35a)$$

$$\Delta G(x)'' = \frac{1 - p_1}{\sigma_2^2 \sqrt{2\pi}} x \exp\left(-\frac{1}{2} \left(\frac{x}{\sigma_2}\right)^2\right) . \quad (4.35b)$$

When $x_{lp} \leq x < 0$, $\Delta G(x)''$ is negative over the domain, indicating that $\Delta G(x)$ is a concave function. We further examine the sign of $\Delta G(x)$ at the two endpoints as follows:

$$\Delta G(x_{lp}) = 0 \quad (4.36a)$$

$$\Delta G(0) = G_o^{\text{core}}(0) - G(0) = 0 . \quad (4.36b)$$

According to the definition of the concave function, for any $\omega \in [0, 1]$, the following equation holds:

$$\Delta G((1 - \omega)x_{lp} + \omega \times 0) \geq (1 - \omega) \Delta G(x_{lp}) + \omega \Delta G(0) = 0 . \quad (4.37)$$

It is equivalent to saying that

$$\Delta G(x) \geq 0 \quad \forall x_{lp} \leq x < 0. \quad (4.38)$$

Similarly, when $0 < x \leq x_{rp}$, $\Delta G(x)''$ is positive over the domain, which means that $\Delta G(x)$ is a convex function. The sign of $\Delta G(x)$ at x_{rp} is examined as follows:

$$\Delta G(x_{rp}) c(x_{rp} - x_{lp}) - (1 - p_1) G(x_{rp}; 0, \sigma_2) + (1 - p_1) G(x_{lp}; 0, \sigma_2). \quad (4.39)$$

By substituting Equations (4.23), (4.33) and $G(x_{rp}; 0, \sigma_2) = 1 - G(x_{lp}; 0, \sigma_2)$ into Equation (4.39), we can obtain that $\Delta G(x_{rp}) = 0$. According to the definition of the convex function, we have

$$\Delta G(x) \leq 0 \quad \forall 0 \leq x \leq x_{lp}. \quad (4.40)$$

Combining Equations (4.38) and (4.40), we can conclude that $G_o^{\text{core}}(x)$ is the CDF overbound of $G(x)$ at the core region.

The PGO of the BGMM with $p_1 = 0.9$, $\sigma_1 = 0.5$, and $\sigma_2 = 1.5$ is plotted in Figure 4.2. In this example, the parameters of the PGO are computed as $k = 0.5881$ and $c = 0.0245$. In addition, the two-step Gaussian overbound [65] is also depicted for comparison. The plot reveals that the PDF and CDF of the PGO are closely aligned with those of BGMM in both the tail and core regions, compared to those of the two-step Gaussian overbound. The thumbnail in Figure 4.2a illustrates the distribution of the PGO in the vicinity of the core-tail transition point, where the PDF of the PGO is not continuous at the transition point. Although the leap of PDF at the transition point is not negligible (calculated to be 0.06 through Equation (4.24)), the CDF of the PGO appears exceptionally smooth near the core-tail transition point, as displayed in the thumbnail of Figure 4.2b.

4.3.3.3 Sigma inflation

Equation (4.24) gives the overbound of the zero-mean BGMM distribution; however, in the application of bounding arbitrary error distributions, samples may not be well

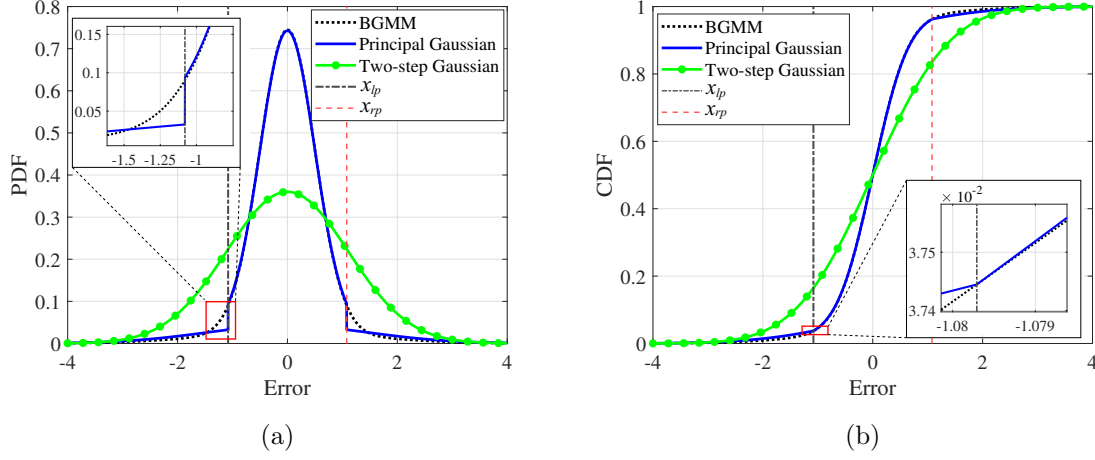


Figure 4.2: The (a) PDF and (b) CDF of the Principal Gaussian overbound of a zero-mean bimodal Gaussian mixture model with $p_1 = 0.9$, $\sigma_1 = 0.5$, and $\sigma_2 = 1.5$. The two-step Gaussian overbound is plotted for comparison.

characterized by the zero-mean BGMM distribution. In such a case, the PGO may not provide an overbound for these samples. These unbounded samples usually occur in the tails of the error distribution. This is because samples in the central region usually exhibit a higher likelihood, and therefore the EM algorithm would prioritize the fitting performance of these central-region samples. In this section, we propose to inflate the tail of the PGO to tackle these unbounded samples.

The most straightforward approach is to increase σ_2 . In the meanwhile, we have to ensure that the inflated PGO is the overbound of the before-inflation PGO and, thus, fitted zero-mean BGMM. Define σ_2^* as the inflated σ_2 , and we have

$$\sigma_2^* = \tau_2 \sigma_2, \quad (4.41)$$

where $\tau_2 > 1$ is the tail inflation factor. Then, the inflated tail bound can be written by

$$G_o^{L*}(x) = (1 + k^*) (1 - p_1) G(x; 0, \sigma_2^*) \quad \forall x < x_{lp} \quad (4.42a)$$

$$k^* = \frac{p_1 G(x_{lp}; 0, \sigma_1)}{(1 - p_1) G(x_{lp}; 0, \sigma_2^*)}, \quad (4.42b)$$

where k^* is the new scaling parameter. Since $\sigma_2^* > \sigma_2$, we have $G(x_{lp}; 0, \sigma_2^*) > G(x_{lp}; 0, \sigma_2)$ and thus $k^* < k$. Therefore, it is difficult to compare the magnitude of $G_o^{L*}(x)$ and $G_o^L(x)$. A naive solution is to make $k^* = k$ by scaling σ_1 to σ_1^* , as shown below

$$\frac{p_1 G(x_{lp}; 0, \sigma_1^*)}{(1 - p_1) G(x_{lp}; 0, \sigma_2^*)} = \frac{p_1 G(x_{lp}; 0, \sigma_1)}{(1 - p_1) G(x_{lp}; 0, \sigma_2)}. \quad (4.43)$$

Indeed, Equation (4.43) is satisfied only when σ_1^* is larger than σ_1 .

We further check the bounding conditions in the core region by examining Equation (4.32c) when $x < 0$. The first term $p_1 G(x; 0, \sigma_1)$ and the third term $(1 - p_1) G(x_{lp}; 0, \sigma_2)$ in Equation (4.32c) will increase with the inflation in σ_1 and σ_2 . The second term $c(x - x_{lp})$ in Equation (4.32c) can be re-written by

$$c(x - x_{lp}) = \frac{(1 - p_1) (G(x_{lp}; 0, \sigma_2) - 0.5)(x - x_{lp})}{x_{lp}}, \quad (4.44)$$

where $x_{lp} \leq x < 0$ and $1 - p_1 > 0$. The inflation in σ_2 will increase the value of $G(x; 0, \sigma_2)$, thus enlarging $c(x - x_{lp})$. Therefore, $G_o^{\text{core}*}(x)$ with inflated σ_2 is larger than $G_o^{\text{core}}(x)$. In summary, the inflated PGO has the following property when $x < 0$,

$$\begin{cases} G_o^{L*}(x) > G_o^L(x) & \forall x < x_{lp} \\ G_o^{\text{core}*}(x) > G_o^{\text{core}}(x) & \forall x_{lp} \leq x < 0, \end{cases} \quad (4.45)$$

indicating that the inflated PGO is the overbound of the before-inflation PGO and, thus, the fitted BGMM.

In addition, unbounded samples may also occur in the core region, although such cases are rare. A slightly different inflation strategy could be applied. We only inflate the core of the PGO to bound these samples. Define σ_1^* as the inflated σ_1 , and we have

$$\sigma_1^* = \tau_1 \sigma_1, \quad (4.46)$$

where $\tau_1 > 1$ is the core inflation factor. The inflation of σ_1 only affects the value of $G(x; 0, \sigma_1)$ in $G_o^{\text{core}}(x)$ and k in $G_o^L(x)$. Actually, both $G(x; 0, \sigma_1)$ and k are increased by inflating σ_1 , thereby enlarging the value of $G_o^{\text{core}}(x)$ and $G_o^L(x)$ when $x < 0$. As

a result, the same conclusion as in Equation (4.45) can be drawn. Our inflation strategies in both tail and core regions can ensure the preservation of overbounding properties with respect to the fitted BGMM.

The inflation of σ_2 and σ_1 can be realized by an iterative approach. In each iteration, we inflate σ_2 or σ_1 according to the violation of bounding conditions with a small and fixed inflation factor. Based on the inflated PGO, all samples are examined for the violation of bounding conditions. The iteration process will stop once all samples are CDF overbounded by the inflated PGO. The pseudocode of the sigma inflation strategy is given in Algorithm 1. In addition, Algorithm 2 summarizes the steps for implementing PGO.

Algorithm 1 Sigma Inflation

Input:

Empirical CDF of Samples: $G_n(x)$

Parameters of the before-inflation PGO: $p_1, \sigma_1, \sigma_2, x_{lp}, x_{rp}$

Output:

Inflated sigma: σ_1^*, σ_2^*

- 1: Initialize the inflation factor: $\tau_1 = 1.01, \tau_2 = 1.01$
 - 2: Core condition: $\Xi_{\text{core}} \leftarrow G_o^{\text{core}}(x) \geq G_n(x) \forall x_{lp} < x < 0$ AND $G_o^{\text{core}}(x) \leq G_n(x) \forall 0 < x < x_{rp}$
 - 3: Tail condition: $\Xi_{\text{tail}} \leftarrow G_o^{\text{L}}(x) \geq G_n(x) \forall x \leq x_{lp}$ AND $G_o^{\text{R}}(x) \leq G_n(x) \forall x \geq x_{rp}$
 - 4: **while** Ξ_{core} is false OR Ξ_{tail} is false **do**
 - 5: **if** Ξ_{core} is false **then**
 - 6: $\sigma_1 \leftarrow \tau_1 * \sigma_1$
 - 7: **end if**
 - 8: Update the tail condition Ξ_{tail} with the latest PGO parameter
 - 9: **if** Ξ_{tail} is false **then**
 - 10: $\sigma_2 \leftarrow \tau_2 * \sigma_2$
 - 11: $\sigma_1 \leftarrow \text{Equation (4.43)}$
 - 12: **end if**
 - 13: Update the core condition Ξ_{core} with the latest PGO parameter
 - 14: Update the tail condition Ξ_{tail} with the latest PGO parameter
 - 15: **end while**
 - 16: $\sigma_1^* \leftarrow \sigma_1, \sigma_2^* \leftarrow \sigma_2$
-

Algorithm 2 Implementation of Principal Gaussian Overbound

Input:

Error data: $X = \{x_1, x_2, \dots, x_n\}$

Partition parameter: α
Output:

Parameters of PGO: $p_1, \sigma_1^*, \sigma_2^*, x_{lp}, x_{rp}$

1: Fit a BGMM

 $f(x) = p_1 f_N(x; 0, \sigma_1) + (1 - p_1) f_N(x; 0, \sigma_2)$
 $\leftarrow$ Employ EM algorithm with X

2: $x_{lp}, x_{rp} \leftarrow$ Dominance partition by Equations (4.22), (4.23)

3: $k \leftarrow$ Calculate the scaling parameter by Equation (4.28)

4: $c \leftarrow$ Calculate the shifting parameter by Equation (4.33)

5: $\sigma_1^*, \sigma_2^* \leftarrow$ Algorithm 1 with $p_1, \sigma_1, \sigma_2, x_{lp}, x_{rp}$

4.3.4 Preservation of overbounding property

In GNSS positioning, it is essential to project the range-domain error to the position-domain error as it allows us to estimate the accuracy of the positioning results obtained from GNSS measurements. In this process, the convolution operation is conducted. In this section, we use DeCleene's theorem [63] to prove that the overbounding property of the PGO can be preserved through convolution.

Given that $G_{oa}(x)$ and $G_{ov}(x)$ are the overbound distribution of the error distribution $G_a(x)$ and $G_v(x)$, respectively, it is essential for the overbound distribution to have the following property:

$$G_{oa+ov}(x) \text{ overbound } G_{a+v}(x), \quad (4.47)$$

where

$$G_{oa+ov}(x) = \int_{-\infty}^x f_{oa}(x) * f_{ov}(x) dx \quad (4.48a)$$

$$G_{a+v}(x) = \int_{-\infty}^x f_a(x) * f_v(x) dx \quad (4.48b)$$

and $*$ denotes the convolution operation.

DeCleene proves that the above property is established if $G_{oa}(x)$, $G_{ov}(x)$, $G_a(x)$, and $G_v(x)$ are all unimodal and symmetric distributions [63].

In this work, the zero-mean BGMM in Equation (4.11) and the PGO in Equation (4.24) are inherently symmetric distributions. Therefore, we only need to prove the unimodality of these distributions. The first derivative of the PDF of the zero-mean BGMM is given by

$$f'(x) = -\frac{x}{\sigma_1^2} \cdot p_1 f_N(x; 0, \sigma_1) - \frac{x}{\sigma_2^2} \cdot (1 - p_1) f_N(x; 0, \sigma_2) . \quad (4.49)$$

Clearly,

$$f'(x) > 0 \ \forall x < 0, \ f'(0) = 0, \ f'(x) < 0 \ \forall x > 0 . \quad (4.50)$$

Therefore, the zero-mean BGMM is a unimodal distribution.

For the PGO, Appendix A.5 proves that $f_o(x)$ is a monotonically increasing function when $x < 0$. According to the symmetric property of $f_o(x)$, we can conclude that $f_o(x)$ is a monotonically decreasing function when $x \geq 0$. Therefore, $f_o(x)$ is a unimodal function. This ends the proof that the overbounding property of the PGO is preserved through convolution.

4.3.5 Numerical consideration of convolution

4.3.5.1 Fourier transform

Equation (4.48a) solves the distribution of the linear combination of random variables oa and ov through convolution. Since the two random variables represent the PGO, which is defined as a piecewise function in Equation (4.24), the convolution operation can be a challenging task due to the complexity of the function. However, the Fourier transform (FT) [79] provides an alternative way to compute the convolution. Consider a linear combination of a set of independent random variables $Y_1, Y_2, \dots, Y_n$ as follows:

$$Y = \sum_{i=1}^n a_i Y_i \quad (4.51)$$

where $a_i, i = 1, 2, \dots, n$ is the coefficient. The random variable Y_i could be the PGO or other symmetric overbounds with a zero mean, such as the zero-mean Gaussian

overbound. The distribution of Y is given by

$$f_Y(x) = f_{\tilde{Y}_1}(x) * f_{\tilde{Y}_2}(x) * \dots * f_{\tilde{Y}_n}(x), \quad (4.52)$$

where

$$f_{\tilde{Y}_i}(x) = \frac{1}{|a_i|} f_{Y_i}\left(\frac{x}{|a_i|}\right) \quad (4.53)$$

and $f_{Y_i}(x)$ is the PDF of the random variable Y_i . The proof is detailed in Appendix A.6. According to the convolution theorem, the FT of $f_Y(x)$ can be expressed by

$$\mathcal{F}(f_Y(x)) = \mathcal{F}(f_{\tilde{Y}_1}(x)) \cdot \mathcal{F}(f_{\tilde{Y}_2}(x)) \cdot \dots \cdot \mathcal{F}(f_{\tilde{Y}_n}(x)), \quad (4.54)$$

where $\mathcal{F}(\cdot)$ is the Fourier transform operator and $\cdot$ denotes the point-wise multiplication. The distribution of $f_Y(x)$ is recovered by the inverse Fourier transform (IFT) as follows:

$$f_Y(x) = \mathcal{F}^{-1}\left(\mathcal{F}(f_{\tilde{Y}_1}(x)) \cdot \mathcal{F}(f_{\tilde{Y}_2}(x)) \cdot \dots \cdot \mathcal{F}(f_{\tilde{Y}_n}(x))\right). \quad (4.55)$$

In practice, the FT and IFT are realized by discrete Fourier transform (DFT) and its inverse (IDFT) [79], respectively, which means the PDF $f_{\tilde{Y}_i}(x)$ should be discretized. According to DeCleene's theorem [63], the discrete model should be the overbound for the continuous distribution it replaces. However, the discretization strategy that directly samples the PDF at equal intervals cannot guarantee the preservation of overbounding properties. Therefore, we propose an alternative discretization strategy, as illustrated in the next section.

4.3.5.2 Discretization satisfying overbounding

Inspired by the discrete overbounding model proposed in [80], we propose to discretize the CDF related to $f_{\tilde{Y}_i}(x)$, which can be formalized as follows:

$$\begin{aligned} F_{\tilde{Y}_i}(x) &= \int_{-\infty}^x f_{\tilde{Y}_i}(x) dx = \int_{-\infty}^x \frac{1}{|a_i|} f_{Y_i}\left(\frac{x}{|a_i|}\right) dx \\ &= F_{Y_i}\left(\frac{x}{|s_{3,i}|}\right), \end{aligned} \quad (4.56)$$

where $F_{Y_i}(x)$ is the CDF related to $f_{Y_i}(x)$. The discretization process can be regarded as constructing a discretized overbound distribution for $F_{\tilde{Y}_i}(x)$. Specifically, Let $2L-1$ be the length of the discretized sequence, and T be the sampling interval (unit: meter), and then we can (almost) equally divide the domain of $F_{\tilde{Y}_i}(x)$ into $2L$ intervals, including $(-\infty, x_1)$, $[x_1, x_2)$, ..., $[x_{2L-2}, x_{2L-1})$, and $[x_{2L-1}, \infty)$, where $x_k = (k - L)T$, $k = 1, 2, \dots, 2L - 1$. The discretization results are shown in Figure 4.3a, where each interval has equal length T except the first and the last interval. The discrete CDF overbound of $F_{\tilde{Y}_i}(x)$ is formalized as a piecewise function as follows:

$$F_{D, \tilde{Y}_i}(x) = \begin{cases} F_{\tilde{Y}_i}(x) & \text{if } x < x_1 \\ F_{\tilde{Y}_i}(x_{k+1}) & \text{if } x_k \leq x < x_{k+1}, 1 \leq k < L - 1 \\ F_{\tilde{Y}_i}(x_k) & \text{if } x_k \leq x < x_{k+1}, L - 1 \leq k < 2L - 2 \\ F_{\tilde{Y}_i}(x) & \text{if } x \geq x_{2L-1} \end{cases}. \quad (4.57)$$

As can be seen, the discrete model $F_{D, \tilde{Y}_i}(x)$ is the overbound for the continuous model $F_{\tilde{Y}_i}(x)$. The probability mass function (PMF) of $F_{D, \tilde{Y}_i}(x)$ can be calculated by

$$p_{D, \tilde{Y}_i}(x) = \begin{cases} F_{D, \tilde{Y}_i}(x_1) - F_{D, \tilde{Y}_i}(x_1) & \text{if } x = x_1 \\ F_{D, \tilde{Y}_i}(x_k) - F_{D, \tilde{Y}_i}(x_{k-1}) & \text{if } x = x_k, 1 < k \leq 2L - 1 \\ 0 & \text{otherwise} \end{cases}, \quad (4.58)$$

which has a discrete nature. The discretization process is completed by evaluating $p_{D, \tilde{Y}_i}(x)$ at the $2L - 1$ points as follows:

$$D_{\tilde{Y}_i}[k] = p_{D, \tilde{Y}_i}(x_k), \quad k = 1, 2, \dots, 2L - 1. \quad (4.59)$$

Then, the PMF of Y in the discretized form can be obtained by the discrete convolution as follows:

$$D_Y[k] = (D_{\tilde{Y}_1} * D_{\tilde{Y}_2} * \dots * D_{\tilde{Y}_n})[k]. \quad (4.60)$$

By taking the discrete convolution theorem, $D_Y[k]$ can be computed as

$$D_Y[k] = \mathcal{F}_D^{-1} \left(\mathcal{F}_D(D_{\tilde{Y}_1}) \cdot \mathcal{F}_D(D_{\tilde{Y}_2}) \cdot \dots \cdot \mathcal{F}_D(D_{\tilde{Y}_n}) \right), \quad k = 1, 2, \dots, L_o \quad (4.61)$$

$$L_o = k \times (2L - 2) + 1,$$

where $\mathcal{F}_D(\cdot)$ and $\mathcal{F}_D^{-1}(\cdot)$ denote the DFT and IDFT [79], respectively. Note that the length of $D_Y[k]$ is extended to $n \times (2L - 2) + 1$. This is because the length of the resulting sequence in the convolution is given by $L_1 + L_2 - 1$, where L_1 and L_2 are the lengths of two input sequences, respectively. In addition, the convolution process does not change the sampling interval; therefore, the distance of the domain of any two adjacent elements in $D_Y[k]$ is T , which is the same as that in $D_{\tilde{Y}_i}[k]$. The DFT is usually implemented by the fast Fourier transform algorithm [81] in the modern software solution or even dedicated hardware, whose computational complexity is only $O((L_e)\log(L_e))$, where L_e is the length of the input sequence and $L_e = 2L - 1$ in our case.

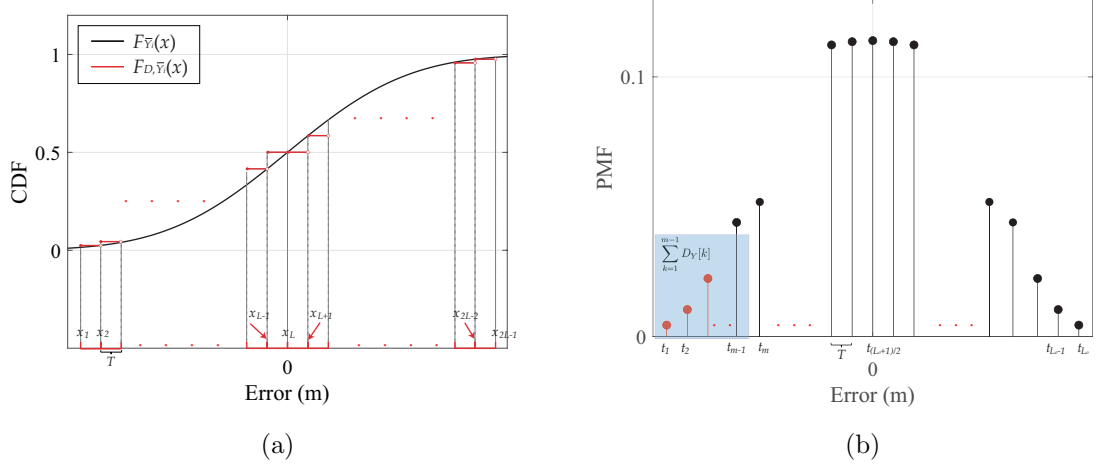


Figure 4.3: Demonstrations of (a) the discretization process and (b) the quantile searching process. The shaded area in (b) shows the cumulative probability from t_1 to t_{m-1} .

4.3.5.3 Finding the Quantile

In integrity monitoring research, finding the quantile of the distribution of test statistics or the positioning error is essential to determine the detection threshold or the protection level. These distributions are usually the results of convolutions of overbounds of measurement errors, which can be calculated by methods developed in Sections 4.3.5.1 and 4.3.5.2. This section shows how to calculate the quantile of the resultant discrete distribution in Equation (4.61). Specifically, the discrete sequence $D_Y[k]$ in Equation (4.61) can be interpreted as the PMF evaluated at $t_k = (k - \frac{L_o + 1}{2})T$. Here we use a different notation t to represent the domain of $D_Y[k]$ without loss of generality. The quantile function $Q_Y^{-1}(\alpha)$ with $0 < \alpha < 1$ can be obtained by finding the index $m \in \{1, \dots, L_o\}$ that satisfies the following conditions:

$$\begin{aligned} \sum_{k=1}^{m-1} D_Y[k] &< \alpha \\ \sum_{k=1}^m D_Y[k] &\geq \alpha, \end{aligned} \tag{4.62}$$

where $\sum_{k=1}^{m-1} D_Y[k]$ is the cumulative probability from t_1 to t_{m-1} , as shown in Figure 4.3b. The quantile function is given by

$$Q_Y^{-1}(\alpha) = t_{m-1} = \left((m-1) - \frac{L_o + 1}{2} \right) T. \tag{4.63}$$

4.4 Bounding Performance of SISRE

In Section 3.4.1 and Section 3.4.2, we depict the QQ plot of $SISRE_{UPE}$ for each GPS and Galileo satellite, where significant differences are observed among satellites. Three categories of $SISRE_{UPE}$ distributions have been identified as follows: 1) Two-side heavy-tailed $SISRE_{UPE}$; 2) One-side heavy-tailed $SISRE_{UPE}$; and 3) Gaussian-like $SISRE_{UPE}$. In this section, we analyze the bounding performance of the proposed PGO on these three types of $SISRE_{UPE}$ distributions. For each error

type, we select one typical satellite from each constellation for detailed analysis. The proposed method is compared with the following two methods:

1. Gaussian overbound: the overbound method used in baseline ARAIM [20, 63];
2. Gaussian-Pareto overbound: a recently developed core overbounding method that bounds the core part of the error with Gaussian distribution and bounds the tail part with the generalized Pareto distribution [68].

4.4.1 Two-side heavy-tailed SISRE

In this section, GPS satellite SVN63 and Galileo satellite GSAT0206 are chosen for analysis. Figure 4.4a shows the folded CDF of $SISRE_{UPE}$ for GPS satellite SVN63, which exhibits significant heavy tails on both sides. The folded CDF of the Gaussian overbound, Gaussian-Pareto overbound, and the proposed PGO are plotted in Figure 4.4a, and the fitted BGMM of the $SISRE_{UPE}$ is also plotted for comparison. As can be seen, the proposed PGO exhibits a tighter bound than the Gaussian overbound in both the core and tail regions of the error distribution. The PGO exhibits a slightly tighter bound than the Gaussian-Pareto overbound in the core region. However, the Gaussian-Pareto overbound shows the sharpest bound at the left tail region. This is because the Gaussian-Pareto overbounding method divides the samples into two individual parts (core and tail parts) and bounds each part separately. In the tail part bounding, the Gaussian-Pareto overbounding method uses the generalized Pareto distribution, which is especially suitable for modeling extreme tails that extend beyond the range of available data [68]. The PGO only shows moderate bounding performance at the left tail region. However, the PGO has its own advantages in that its overbounding property can be preserved through convolution, which is essential in the range-to-position projection process and the calculation of PLs. It is worth noting that the fitted BGMM fails to overbound the error distribution, as shown in the regions of $x \in [-15 \text{ m}, -2 \text{ m}]$ and $x \in [2 \text{ m}, 15 \text{ m}]$ in Figure 4.4a. This is because

the fitted BGMM can only fit the overall pattern of the error distribution and is not guaranteed to bound the error distribution.

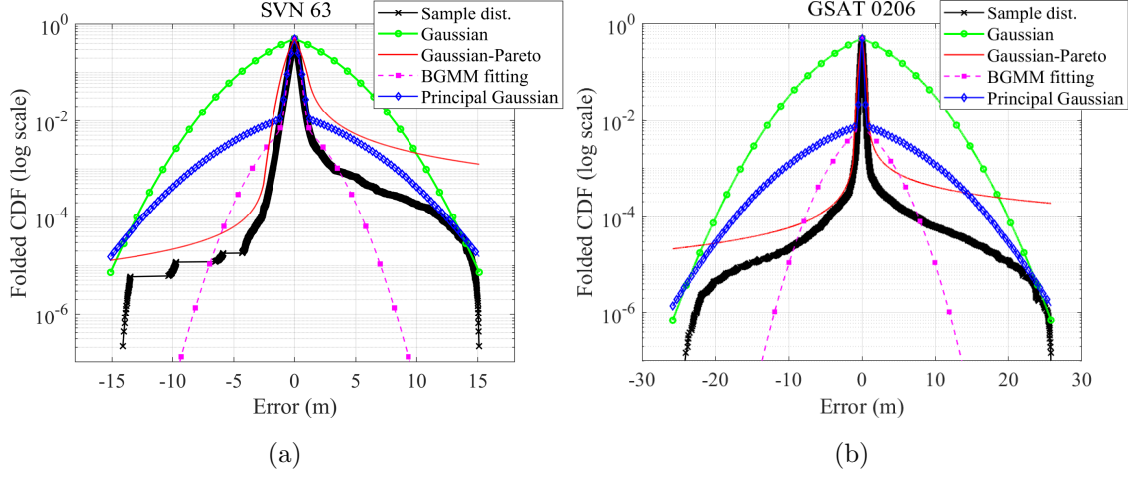


Figure 4.4: The folded CDF of $SISRE_{UE}$ and its bounding results for (a) GPS satellite SVN63; and (b) Galileo satellite GSAT0206.

Figure 4.4b shows the folded CDF of $SISRE_{UE}$ for Galileo satellite GSAT0206. Similar to GPS satellite SVN63, GSAT0206 has a two-side heavy-tailed $SISRE_{UE}$. However, the $SISRE_{UE}$ of Galileo satellite GSAT0206 has a narrower core than that of GPS satellite SVN63. In addition, the maximum absolute $SISRE_{UE}$ reaches 26 m, which is around 73 % larger than that observed in the SVN63's $SISRE_{UE}$ (15 m in this case). This finding suggests that the $SISRE_{UE}$ of GSAT0206 has heavier tails than that of SVN63. In this condition, the Gaussian-Pareto overbound has the best performance in bounding the core region. However, the Gaussian-Pareto overbound experiences a significant performance degradation when the folded CDF is down to 10^{-4} in the left tail region and 10^{-3} in the right tail region. The proposed PGO shows moderate bounding performance at the core region, which has a tighter bound than the Gaussian overbound but not as tight as the Gaussian-Pareto overbound. In the tail region, the proposed PGO keeps producing tighter bounds than the Gaussian overbound. In summary, the proposed PGO shows the most competitive performance in bounding $SISRE_{UE}$ of Galileo satellite GSAT0206.

4.4.2 One-side heavy-tailed SISRE

In this section, GPS satellite SVN66 and Galileo satellite GSAT0212 are chosen for analysis. Figures 4.5a and 4.5b depict the $SISRE_{UPE}$ distribution of these two satellites in the folded CDF view, respectively. SVN66's $SISRE_{UPE}$ has a left-side heavy tail, while GSAT0212's $SISRE_{UPE}$ has a right-side heavy tail. Though different from the two-side heavy-tailed case in Section 4.4.1, the $SISRE_{UPE}$ of the Galileo satellite still has a narrower core than that of the GPS satellite. Since the Gaussian overbound and the proposed PGO have a symmetric form, their tail-bounding results are mostly determined by the errors on the heavy-tailed side. As can be seen, the PGO and the Gaussian overbound closely align with the right-tail errors in Figure 4.5a and the left-tail errors in Figure 4.5b. As a result, the PGO and the Gaussian overbound exhibit loose bounds on the light-tailed side. Nevertheless, the PGO consistently yields a tighter bound than the Gaussian overbound at all error values in both Figure 4.5a and Figure 4.5b.

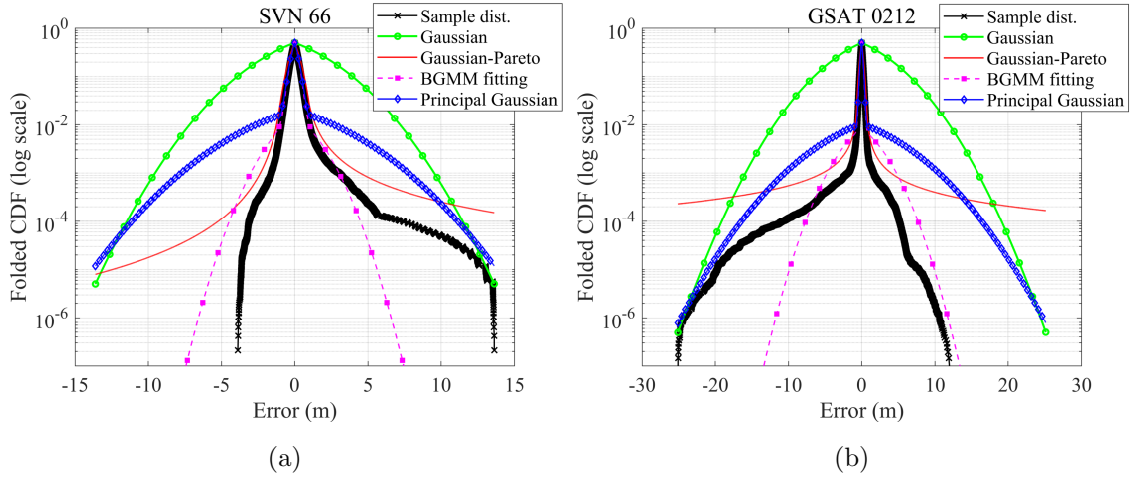


Figure 4.5: The folded CDF of $SISRE_{UPE}$ and its bounding results for (a) GPS satellite SVN66; and (b) Galileo satellite GSAT0212.

The Gaussian-Pareto overbound is not limited to a symmetric form, and therefore its bounding results are determined by the errors on each side separately, producing

asymmetry bounding results. While the Gaussian-Pareto overbound provides tight bounds in the core region, it is intentionally conservative in the far tail regions to account for the uncertainty envelope around the empirical distribution [68]. In addition, the asymmetry in Gaussian-Pareto overbounds can cause difficulties in preserving the overbounding properties after convolution, making it impossible to calculate the PL. This, on the other hand, highlights the benefits of applying the proposed method to bound the heavy-tailed distribution, as it enables the computation of PL for integrity-assured applications.

4.4.3 Gaussian-like SISRE

In this section, only the GPS satellite is chosen for analysis, as there is no Gaussian-like $SISRE_{UPE}$ in the Galileo satellites. Specifically, we choose $SISRE_{UPE}$ of GPS satellite SVN46 for analysis. Figure 4.6 shows the distribution of $SISRE_{UPE}$ and the bounding results of three methods. As can be seen, the error distribution is not significantly heavy-tailed, which is also suggested by the QQ plot in Figure 3.8. In this situation, there seems to be no significant difference among the three overbounding methods. Therefore, it is recommended to use the Gaussian overbound when the error distribution does not exhibit heavy-tailed properties, as the Gaussian overbound has fewer parameters to determine.

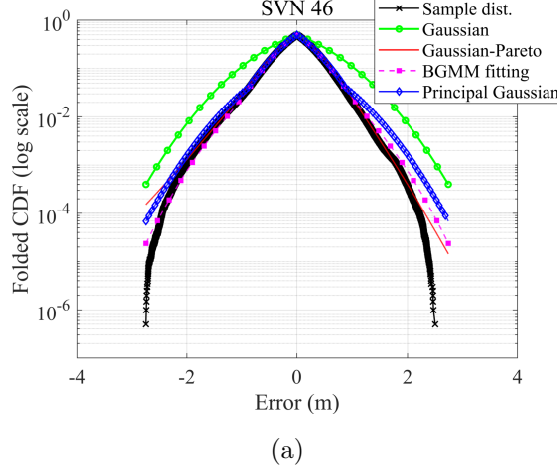


Figure 4.6: The folded CDF of $SISRE_{UE}$ and its bounding results for GPS satellite SVN46.

4.4.4 Bounding parameters for individual satellites

Tables 4.1 and 4.2 give the bounding parameters of the Gaussian overbound and the PGO for GPS and Galileo SISRE, respectively. The 1-sigma Gaussian overbound of GPS SISRE varies significantly, with an average of 1.67 m. This is because the SISRE of some GPS satellites exhibits heavy-tailed properties while the others have Gaussian-like behavior, as revealed in Section 3.4.1. This difference is also reflected in the PGO parameters, where the heavy-tailed SISRE featured with a large gap between σ_1 and σ_2 , and the Gaussian-liked SISRE has a smaller deviation between σ_1 and σ_2 .

For the Galileo satellites, the 1-sigma Gaussian overbound of SISRE has a smaller variation, with an average of 5.58 m. This value aligns closely with the Galileo broadcast URA parameter, $\sigma_{URA} = 6$ m, as defined in Galileo OS-SDD [62]. Since the SISRE of all Galileo satellites exhibits significant heavy-tailed properties, the Galileo broadcast URA parameter is likely to provide an extremely conservative bound for the SISRE. The impacts of such conservatism will be further discussed in Chapter 5 and

Chapter 6. For the PGO parameters, all Galileo satellites exhibit a high consistency, where σ_2 is significantly larger than σ_1 , and the p_1 is larger than 0.98.

4.5 Non-Gaussian Nominal Error Bounds of Code IF Combination

The measurement error of the code IF combination is given in Equation (2.11), which consists of the range projection of clock and orbit error, tropospheric error, and multipath and code noise. The nominal error bound of the code IF combination regarding the satellite i and the receiver j can be formalized as follows:

$$G_{\varrho,j,IF,acc}^i(x) = \int_{-\infty}^x f_{\varrho,j,IF,acc}^i(x) dx \quad (4.64a)$$

$$f_{\varrho,j,IF,acc}^i(x) = f_{orb\&clk,PGO}^i(x) * f_{tropo,j,ob}^i(x) * f_{\varrho,user,j,IF,ob}^i(x), \quad (4.64b)$$

where $f_{orb\&clk,PGO}^i(x)$ is the PGO of the range projection of clock and orbit error, whose parameters are given in Tables 3.3 and 3.4; $f_{tropo,j,ob}^i(x)$ is the Gaussian overbound of the tropospheric error with a zero mean and the standard deviation defined in Equation (2.14); $f_{\varrho,user,j,IF,ob}^i(x)$ is the Gaussian overbound of the multipath and code noise with a zero mean and the standard deviation defined in Appendix A.1.

The nominal error bound in Equation (4.64) is developed for accuracy evaluation and fault detection purposes (as will be exploited in Chapter 5), which does not consider error sources that introduce biases in the distributions. However, to protect the integrity, the bound also needs to consider nominal signal deformation errors. These events usually alter the range projection of orbit and clock error distribution, moving its median away from the origin. For integrity purposes, the pair overbounding method [64] introduced the b_{nom} to establish a symmetric error envelope, which is formed by two equally shifted Gaussian distributions. The nominal error bound for accuracy in Equation (4.64) could be obtained by simply integrating the same

principle as follows:

$$G_{\varrho,j,IF,int}^i(x) = \begin{cases} \int_{-\infty}^x f_{\varrho,j,IF,acc}^i(x + b_{nom,i}) \, dx & \text{if } G_v(x) < \frac{1}{2} \\ \frac{1}{2} & \text{otherwise} \\ \int_{-\infty}^x f_{\varrho,j,IF,acc}^i(x - b_{nom,i}) \, dx & \text{if } G_v(x) > \frac{1}{2} \end{cases}, \quad (4.65)$$

where $G_v(x)$ is the empirical distribution of measurement errors of the code IF combination. In [82], b_{nom} is recommended to take 0.75 m to conservatively bound the impacts of nominal signal deformations. The above modification simply results in broader margins for the error distributions to account for signal distortion events. According to [64], the overbounding properties of the bound in Equation (4.65) can still be preserved through convolution. Now, the nominal error bound in Equation (4.65) is developed for integrity, which will be exploited in Chapter 6.

Table 4.1: Parameters of the Gaussian overbound and the Principal Gaussian overbound of $SISRE_{UPE}$ for each GPS satellite (unit: m)

SVN	Type ¹	Gaussian	PGO			
		σ	σ_1	σ_2	p_1	x_{rp}
SVN41	O	1.136	0.403	1.343	0.918	0.948
SVN43	T	1.113	0.432	1.195	0.762	0.906
SVN44	T	4.052	0.595	4.425	0.628	1.103
SVN45	O	1.778	0.425	2.226	0.955	1.157
SVN46	G	0.818	0.413	0.780	0.787	0.884
SVN47	G	0.521	0.351	0.612	0.861	0.835
SVN48	G	0.780	0.414	0.804	0.535	0.611
SVN50	G	0.574	0.411	0.691	0.977	1.493
SVN51	O	2.518	0.385	3.211	0.973	1.042
SVN52	G	0.703	0.427	0.765	0.716	0.893
SVN53	O	2.245	0.540	2.426	0.624	1.074
SVN55	G	0.873	0.310	0.998	0.891	0.763
SVN56	G	0.680	0.382	0.815	0.956	1.120
SVN57	O	1.080	0.471	1.333	0.806	0.790
SVN58	O	2.998	0.372	3.998	0.983	1.136
SVN59	G	0.616	0.297	0.544	0.783	0.667
SVN61	T	0.753	0.321	0.837	0.788	0.684
SVN62	O	0.694	0.355	0.835	0.961	1.046
SVN63	T	3.487	0.419	4.425	0.970	1.073
SVN64	O	1.495	0.390	2.050	0.985	1.155
SVN65	T	3.570	0.353	3.901	0.574	0.669
SVN66	O	3.084	0.363	3.968	0.970	0.963
SVN67	G	0.540	0.292	0.600	0.840	0.649
SVN68	T	0.977	0.302	1.170	0.928	0.707
SVN69	T	3.302	0.468	3.908	0.894	1.034
SVN70	T	2.303	0.308	2.959	0.965	0.821
SVN71	T	0.934	0.341	1.112	0.920	0.832
SVN72	G	1.548	1.005	1.441	0.548	0.872
SVN73	T	3.680	0.521	4.110	0.842	1.154
SVN74	O	1.287	0.310	1.602	0.973	0.839

¹ “T”: Two-side heavy-tailed; “O”: One-side heavy-tailed; “G”: Gaussian-liked.

Table 4.2: Parameters of the Gaussian overbound and the Principal Gaussian overbound of $SISRE_{UPE}$ for each Galileo satellite (unit: m)

SVN	Type ¹	Gaussian	PGO			
		σ	σ_1	σ_2	p_1	x_{rp}
GSAT0101	T	5.967	0.292	7.717	0.985	0.790
GSAT0102	T	5.758	0.311	7.662	0.984	0.909
GSAT0103	T	6.098	0.289	7.430	0.980	0.752
GSAT0203	T	5.890	0.338	8.278	0.986	0.967
GSAT0205	O	2.333	0.229	2.867	0.984	0.680
GSAT0206	T	5.346	0.236	6.859	0.986	0.717
GSAT0207	O	5.724	0.256	7.188	0.983	0.758
GSAT0208	T	5.687	0.246	7.144	0.985	0.740
GSAT0209	T	5.423	0.232	7.245	0.986	0.682
GSAT0210	T	5.714	0.230	8.783	0.980	0.570
GSAT0211	O	6.197	0.234	7.809	0.984	0.715
GSAT0212	O	5.136	0.250	6.351	0.983	0.725
GSAT0213	T	5.970	0.251	8.416	0.984	0.691
GSAT0214	T	5.561	0.238	6.926	0.983	0.693
GSAT0215	T	5.619	0.238	7.483	0.985	0.694
GSAT0216	T	7.383	0.229	9.264	0.983	0.698
GSAT0217	T	5.518	0.228	7.160	0.986	0.673
GSAT0218	T	5.598	0.229	7.031	0.983	0.676
GSAT0219	T	6.155	0.280	7.761	0.986	0.795
GSAT0220	T	5.000	0.297	6.404	0.985	0.877
GSAT0221	T	5.266	0.269	6.579	0.986	0.799
GSAT0222	T	5.332	0.259	6.663	0.980	0.723
GSAT0223	O	5.521	0.288	7.300	0.988	0.895
GSAT0224	O	5.644	0.275	7.458	0.987	0.860

¹ “T”: Two-side heavy-tailed; “O”: One-side heavy-tailed.

Chapter 5

Fault Detection with Non-Gaussian Nominal Error Bound

Fault detection is essential for positioning and navigation systems in some safety-critical applications [15, 83–86], which is the technology to check the occurrence of faults in the system as well as to determine the time of fault occurs [87]. Most fault detection methods, such as range comparison method [4], parity space [9, 88], chi-squared test [2, 89], and solution separation [6, 16], assume that the nominal measurement error is Gaussian distributed. However, chapter 3 reveals that the SISRE distribution of GPS and Galileo satellites has heavy tails. The unrealistic Gaussian assumptions can result in degraded fault detection rates in real-world applications, limiting the reliability and effectiveness of preventing systems from faults.

In this chapter, a rigorous hypothesis testing method is developed to detect faulty measurements in navigation systems under non-Gaussian nominal errors by introducing the jackknife technique, a cross-validation technique in statistics [90, 91]. The basic idea is to quantify the inconsistency between the observed measurement and the predicted measurement based on subset solutions. Section 5.1 constructs the jackknife detector for single fault detection under non-Gaussian nominal errors. Section

5.2 extends this technique to multi-fault detection. Section 5.3 demonstrates the excellent performance of the proposed method in a worldwide simulation by comparing it with the solution separation detector.

5.1 Jackknife Detector for Single Fault Detection

In statistics, the jackknife is a cross-validation technique, initially developed by [92] and expanded and named by [90]. The basic idea of the jackknife technique is to systematically leave out each observation from a dataset and calculate the parameter estimate over the remaining observations. Then, these calculations are aggregated for specific statistical purposes [90, 91]. This section shows how to derive the jackknife residual for linearized pseudorange-based positioning systems and develop the hypothesis test to detect potential faulty measurements.

5.1.1 Full set and subset solutions based on weighted least square

The proposed method shares the common logic of solution separation to compute the full set and subset solutions, as shown in Section 2.3.4. To ease the reading, the linearized measurement model defined in Equation (2.20) is re-written here

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \boldsymbol{\varepsilon}. \quad (5.1)$$

With n measurements, the full set solution $\hat{\mathbf{x}}_t$ can be solved by the WLS method, as shown in Equation (2.35):

$$\begin{aligned} \hat{\mathbf{x}} &= \mathbf{S}\mathbf{y} \\ \hat{\mathbf{x}}_t &= \mathbf{x}_0 + \hat{\mathbf{x}}, \end{aligned} \quad (5.2)$$

where

$$\mathbf{S} = (\mathbf{G}^T \mathbf{W} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}. \quad (5.3)$$

To obtain the k th subset solution, the measurements with indices $i \notin idx_k^{ex}$ are excluded. In this section, only single faulty measurement is considered, i.e., $|idx_k^{ex}| = 1$. Then the measurement model in Equation (2.20) can be re-written by

$$\mathbf{y}^{(k,*)} = \mathbf{G}^{(k,*)} \mathbf{x}^{(k)} + \boldsymbol{\varepsilon}^{(k,*)} \quad (5.4)$$

where $\mathbf{y}^{(k,*)}$, $\boldsymbol{\varepsilon}^{(k,*)}$, $\mathbf{x}^{(k)}$ and $\mathbf{G}^{(k,*)}$ have the same meaning as that in Equation (2.20) but are defined on the k th subset. Note that $\mathbf{y}^{(k,*)}$ and $\boldsymbol{\varepsilon}^{(k,*)}$ are $(n-1) \times 1$ vectors, $\mathbf{x}^{(k)}$ is an $m \times 1$ vector, and $\mathbf{G}^{(k,*)}$ is a $(n-1) \times m$ matrix. Define the $m \times (n-1)$ matrix $\mathbf{S}^{(k,*)}$ as the solution matrix on the k th subset

$$\mathbf{S}^{(k,*)} = \left(\mathbf{G}^{(k,*)T} \mathbf{W}^{(k,*)} \mathbf{G}^{(k,*)} \right)^{-1} \mathbf{G}^{(k,*)T} \mathbf{W}^{(k,*)}, \quad (5.5)$$

where $\mathbf{W}^{(k,*)}$ is the weight matrix and is constructed by remove the $i \in idx_k^{ex}$ rows and columns of $\mathbf{W}$. Here, $\mathbf{W}^{(k,*)}$ has the size of $(n-1) \times (n-1)$. An alternative construction of the subset solution matrix is given as follows:

$$\mathbf{S}^{(k)} = (\mathbf{G}^T \mathbf{W}^{(k)} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}^{(k)}, \quad (5.6)$$

where $\mathbf{W}^{(k)}$ is a diagonal matrix and is defined as

$$W_{i,i}^{(k)} = \begin{cases} 0 & \text{if } i = k \\ W_{i,i} & \text{otherwise} \end{cases}. \quad (5.7)$$

Now, $\mathbf{S}^{(k)}$ can operate on the full of measurements. The subsolutions are given by

$$\hat{\mathbf{x}}^{(k)} = \mathbf{S}^{(k)} \mathbf{y} \quad \forall k = 1 \cdots n \quad (5.8a)$$

$$\hat{\mathbf{x}}_t^{(k)} = \mathbf{x}_0 + \hat{\mathbf{x}}^{(k)} \quad \forall k = 1 \cdots n, \quad (5.8b)$$

where $\hat{\mathbf{x}}_t^{(k)}$ is the estimation of the positioning state $\mathbf{x}_t^{(k)}$ associated with the k th subset.

5.1.2 Construction of jackknife residual

The predicted k th measurement with the subsolution $\hat{\mathbf{x}}^{(k)}$ is given by

$$\hat{y}_k = \mathbf{g}_k \hat{\mathbf{x}}^{(k)}, \quad (5.9)$$

where $\mathbf{g}_k$ is the k th row of $\mathbf{G}$. The jackknife residual is given by the difference between y_k and $\hat{y}_k$

$$t_k = y_k - \hat{y}_k, \quad (5.10)$$

where y_k is the k th element of $\mathbf{y}$.

5.1.3 Distribution of jackknife residual

The predicted measurement vector $\hat{\mathbf{y}}^{(k,*)}$ based on the subsolution $\hat{\mathbf{x}}^{(k)}$ is given by

$$\hat{\mathbf{y}}^{(k,*)} = \mathbf{G}^{(k,*)} \hat{\mathbf{x}}^{(k)}, \quad (5.11)$$

where $\hat{\mathbf{y}}^{(k,*)}$ is a $(n-1) \times 1$ column vector and $\mathbf{G}^{(k,*)}$ is a $(n-1) \times m$ matrix. Construct $\tilde{\mathbf{y}}^{(k)}$ as follows:

$$\tilde{\mathbf{y}}^{(k)} = \begin{bmatrix} \hat{\mathbf{y}}_{1:k-1,:}^{(k,*)} \\ \hat{y}_k \\ \hat{\mathbf{y}}_{k:n-1,:}^{(k,*)} \end{bmatrix}, \quad (5.12)$$

where $\hat{\mathbf{y}}_{1:k-1,:}^{(k,*)}$ is the first $k-1$ rows of $\hat{\mathbf{y}}^{(k,*)}$ and $\hat{\mathbf{y}}_{k:n-1,:}^{(k,*)}$ is the remaining rows of $\hat{\mathbf{y}}^{(k,*)}$.

The modified measurement residual is given by

$$\begin{aligned} \mathbf{y} - \tilde{\mathbf{y}}^{(k)} &= \mathbf{y} - \mathbf{G} \hat{\mathbf{x}}^{(k)} \\ &= (\mathbf{I} - \tilde{\mathbf{P}}^{(k)}) \mathbf{y}, \end{aligned} \quad (5.13)$$

where

$$\tilde{\mathbf{P}}^{(k)} = \mathbf{G} \mathbf{S}^{(k)}. \quad (5.14)$$

According to Equations (5.1) and (5.4), $\mathbf{y}$ can be re-written by

$$\mathbf{y} = \mathbf{G} \mathbf{x}^{(k)} + \boldsymbol{\varepsilon}, \quad (5.15)$$

where $\boldsymbol{\varepsilon}$ is the measurement error vector with arbitrary distributions. Then, Equation (5.13) can be written by

$$\mathbf{y} - \tilde{\mathbf{y}}^{(k)} = (\mathbf{I} - \tilde{\mathbf{P}}^{(k)}) \mathbf{G} \mathbf{x}^{(k)} + (\mathbf{I} - \tilde{\mathbf{P}}^{(k)}) \boldsymbol{\varepsilon}. \quad (5.16)$$

Since

$$\begin{aligned}
 (\mathbf{I} - \tilde{\mathbf{P}}^{(k)}) \mathbf{G} &= (\mathbf{I} - \mathbf{G} \mathbf{S}^{(k)}) \mathbf{G} \\
 &= (\mathbf{I} - \mathbf{G} (\mathbf{G}^T \mathbf{W}^{(k)} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W}^{(k)}) \mathbf{G} \\
 &= 0,
 \end{aligned} \tag{5.17}$$

we have

$$\mathbf{y} - \tilde{\mathbf{y}}^{(k)} = (\mathbf{I} - \tilde{\mathbf{P}}^{(k)}) \boldsymbol{\varepsilon}. \tag{5.18}$$

Define $\tilde{\mathbf{p}}_k^{(k)}$ as the k th row of $(\mathbf{I} - \tilde{\mathbf{P}}^{(k)})$, then the jackknife residual is given by

$$t_k = \tilde{\mathbf{p}}_k^{(k)} \boldsymbol{\varepsilon}. \tag{5.19}$$

Equation (5.19) can be rewritten as the linear combination of measurement errors as follows:

$$t_k = \sum_{j=1}^n \tilde{p}_{k,j}^{(k)} \varepsilon_j, \tag{5.20}$$

where $\tilde{p}_{k,j}^{(k)}$ is the j th element of $\tilde{\mathbf{p}}_k^{(k)}$, and ε_j is the j th element of $\boldsymbol{\varepsilon}$. Remarkably, ε_j can have an **arbitrary distribution** as long as it has a PDF $f_{\varepsilon_j}(\cdot)$. Since t_k is the weighted sum of independent random variables with zero-mean distributions, its PDF can be easily obtained by (see Appendix A.6).

$$f_{t_k}(x) = \prod_{j=1}^n |\tilde{p}_{k,j}^{(k)}|^{-1} f_{\varepsilon_1} \left(\frac{x}{|\tilde{p}_{k,1}^{(k)}|} \right) * f_{\varepsilon_2} \left(\frac{x}{|\tilde{p}_{k,2}^{(k)}|} \right) * \dots * f_{\varepsilon_n} \left(\frac{x}{|\tilde{p}_{k,n}^{(k)}|} \right). \tag{5.21}$$

In the special case where ε_j has a zero-mean Gaussian distribution, i.e.,

$$\varepsilon_j \sim \mathcal{N}(0, \sigma_j^2) \quad \forall j = 1 \dots n, \tag{5.22}$$

the distribution of t_k is given by (a proof is provided in Appendix A.7)

$$t_k \sim \mathcal{N} \left(0, \mathbf{g}_k \mathbf{S}^{(k)} \mathbf{W}^{-1} \mathbf{S}^{(k)T} \mathbf{g}_k^T + \sigma_k^2 \right). \tag{5.23}$$

5.1.4 Jackknife test for fault detection

Formalize the following hypotheses:

$$\begin{aligned} H_{0,k}: & \text{No failure in the } k\text{th measurement} \\ H_{1,k}: & \text{A failure in the } k\text{th measurement.} \end{aligned} \tag{5.24}$$

The hypothesis testing for fault detection can be formalized by:

Origin test: $H_{0,k}$ is rejected if $|t_k| > \left(\mathbf{g}_k \mathbf{S}^{(k)} \mathbf{W}^{-1} \mathbf{S}^{(k)T} \mathbf{g}_k^T + \sigma_k^2 \right)^{\frac{1}{2}} Q^{-1}\left(\frac{\alpha}{2}\right)$ at significant level of α , where $Q^{-1}(\cdot)$ is the quantile function of a standard normal variable. The probability of Type I error (false alert) of the origin test is α .

In practice, the above test will be conducted for each subsolution to detect the potential failure in measurements, which evolves into a multiple-testing problem. In such a case, the Type I error is actually increased. Thus, the following hypotheses are formalized instead, which are known as the Bonferroni correction [93]:

$$\begin{aligned} H_0: & \text{No failure in the } n \text{ measurements} \\ H_1: & \text{At least one failure in the } n \text{ measurements.} \end{aligned} \tag{5.25}$$

The hypothesis testing using the corrected hypotheses is formalized by:

Corrected test: H_0 is rejected if $|t_k| > \left(\mathbf{g}_k \mathbf{S}^{(k)} \mathbf{W}^{-1} \mathbf{S}^{(k)T} \mathbf{g}_k^T + \sigma_k^2 \right)^{\frac{1}{2}} Q^{-1}\left(\frac{\tau}{2n}\right)$ at significant level of α^* , where τ is the upper limit of α^* . The probability of Type I error of the corrected test is α^* .

In implementing the corrected test, τ will be specified (e.g., 0.05) according to the nature of the application. Then the probability of Type I error of the individual test would be $\alpha = \frac{\tau}{n}$ (as shown in Appendix A.8), which could be very small when n takes a large value. Therefore, the individual test and the corrected test both could be conservative. However, in satellite navigation applications, it is rare to have a large n , which ensures the feasibility of the corrected test.

5.2 Extend Jackknife Detector to Simultaneous Faults

The developed jackknife detector in Section 5.1 has an underlying assumption that at most one faulty measurement occurs per time epoch. This single-fault assumption was valid in the early stage of satellite navigation with limited satellites in operation [4, 8, 15]. However, with the growing number of satellites and constellations, the probability of simultaneous faults becomes non-negligible. For example, multiple GPS satellites experienced high L1 single-frequency range errors of up to 16 m due to an erroneous ionospheric correction term between May 28 and June 2, 2002 [94]. This highlights the need for fault detection techniques in handling multiple faults [20].

In fact, researchers have already proposed an optimal fault detection algorithm under certain assumptions [95]. This algorithm involves evaluating the consistency of all sets of measurements and selecting the best set with the highest level of consistency. The implementation of this algorithm in the navigation community can refer to the multiple-hypothesis solution separation for multiple faults integrity monitoring [20]. In this section, we leverage this idea to extend the jackknife detector to multiple fault detection with non-Gaussian nominal errors.

5.2.1 Reconstruction of jackknife residual

The threat model defined in [20] is utilized to re-construct the jackknife residual in Equation (5.10) to handle the multiple-fault condition. The threat model defines a collection of error modes that partition the whole measurement space. The fault mode 0 represents the fault-free case while other fault modes (i.e., $1, 2, \dots, N_{\text{fault modes}}$) indicate the presence of single or multiple faults. The exact form of the threat model is detailed in Section 2.3.3 and will not be dwelt upon here.

For fault mode k , the weight matrix in Equation (5.6) can be re-constructed as follows:

$$W_{i,i}^{(k)} = \begin{cases} 0 & \text{if } i \in idx_k^{ex} \\ W_{i,i} & \text{otherwise} \end{cases}. \quad (5.26)$$

The jackknife residual regarding the $i \in idx_k^{ex}$ th measurement for fault mode k is given by

$$t_i^{(k)} = y_i - \hat{y}_i^{(k)}, \quad (5.27)$$

where $\hat{y}_i^{(k)}$ is the predicted i th measurement based on subset solution $\hat{\mathbf{x}}^{(k)}$, as defined in Equation (5.9). It is easy to extend Equation (5.20) to the simultaneous faults condition as follows:

$$t_i^{(k)} = \sum_{j=1}^n \tilde{p}_{i,j}^{(k)} \varepsilon_j, i \in idx_k^{ex}, \quad (5.28)$$

where $\tilde{p}_{i,j}^{(k)}$ is the (i, j) element of $\mathbf{I} - \tilde{\mathbf{P}}^{(k)}$.

It is worth noting that the existence of $t_i^{(k)}$ depends on the existence of the subset solution $\hat{\mathbf{x}}^{(k)}$, which is not guaranteed in the constellation fault mode. This is because all satellite measurements from the hypothetically faulty constellation are excluded in this fault mode, making it impossible to solve the receiver clock bias related to the hypothetically faulty constellation in $\hat{\mathbf{x}}^{(k)}$. Therefore, the constellation fault is temporally not considered in constructing jackknife detectors in the following sections. This problem will be reviewed in Chapter 6.

5.2.2 Combination of jackknife residuals

When $k > n$, there are multiple jackknife residuals associated with fault mode k , making it difficult to construct a hypothesis test. Therefore, the following combination of jackknife residuals is adopted:

$$\tilde{t}_k = \sum_{i \in idx_k^{ex}} S_{v,i} t_i^{(k)}, k = n+1, n+2, \dots, N_{\text{fault modes}}, \quad (5.29)$$

where $S_{v,i}$ is the (v,i) th element of the full set solution matrix $\mathbf{S}$. This kind of weighting scheme can greatly reduce the complexity of developing integrity monitoring algorithms, as will be shown in Chapter 6.

By substituting (5.28) into (5.29), we have

$$\tilde{t}_k = \sum_{j=1}^n \sum_{i \in id x_k^{ex}} S_{v,i} \tilde{p}_{i,j}^{(k)} \varepsilon_j. \quad (5.30)$$

The PDF of $\tilde{t}_k$ can be derived as

$$\begin{aligned} f_{\tilde{t}_k}(x) = \prod_{j=1}^n \left| \sum_{i \in id x_k^{ex}} S_{v,i} \tilde{p}_{i,j}^{(k)} \right|^{-1} & f_{\varepsilon_1} \left(\frac{x}{\left| \sum_{i \in id x_k^{ex}} S_{v,i} \tilde{p}_{i,1}^{(k)} \right|} \right) * f_{\varepsilon_2} \left(\frac{x}{\left| \sum_{i \in id x_k^{ex}} S_{v,i} \tilde{p}_{i,2}^{(k)} \right|} \right) * \\ & \dots * f_{\varepsilon_n} \left(\frac{x}{\left| \sum_{i \in id x_k^{ex}} S_{v,i} \tilde{p}_{i,n}^{(k)} \right|} \right). \end{aligned} \quad (5.31)$$

In the special case of Gaussian noises, i.e., $\varepsilon_j \sim \mathcal{N}(0, \sigma_j^2)$, we have

$$\tilde{t}_k \sim \mathcal{N} \left(0, \sum_{j=1}^n \left(\sum_{i \in id x_k^{ex}} S_{v,i} \tilde{p}_{i,j}^{(k)} \right)^2 \sigma_j^2 \right). \quad (5.32)$$

To unify the notation in the following sections, we define the following test statistics

$$t_k^* = \begin{cases} t_k & \text{if } k = 1, 2, \dots, n \\ \tilde{t}_k & \text{if } k = n+1, n+2, \dots, N_{\text{fault modes}} \end{cases}. \quad (5.33)$$

5.2.3 Reconstruction of hypothesis tests

The following hypotheses are constructed:

$$\begin{aligned} H_0: & \text{The hypothesis corresponding to fault mode 0} \\ H_k: & \text{The hypothesis corresponding to fault mode } k. \end{aligned} \quad (5.34)$$

This, again, involves multiple testing. The reject region for test H_0 v.s. H_k can be defined as

$$R_k = \{t_k^* \mid |t_k^*| \geq T_k\}, k = 1, 2, \dots, N_{\text{fault modes}}, \quad (5.35)$$

where T_k is the threshold for t_k^* . Assume that the probability of the Type I error of the above multiple testing problem is α^* , i.e.,

$$\alpha^* = P\left(\bigcup_{k=1}^{N_{\text{fault modes}}} t_k^* \in R_k \mid H_0\right). \quad (5.36)$$

Since $R_k, k = 1, 2, \dots, N_{\text{fault modes}}$ are not mutually exclusive, we have

$$\begin{aligned} \alpha^* &= P\left(\bigcup_{k=1}^{N_{\text{fault modes}}} t_k^* \in R_k \mid H_0\right) \\ &\leq \sum_{k=1}^{N_{\text{fault modes}}} P(t_k^* \in R_k \mid H_0) \\ &= \sum_{k=1}^{N_{\text{fault modes}}} P(|t_k^*| \geq T_k \mid H_0) = \tau. \end{aligned} \quad (5.37)$$

According to the Bonferroni correction [93], by setting

$$T_k = Q_{t_k^*}^{-1}\left(\frac{\tau}{2N_{\text{fault modes}}}\right), \quad (5.38)$$

H_0 is rejected if any $|t_k^*| > T_k$ at significant level of α^* , where $Q_{t_k^*}^{-1}(\cdot)$ is the quantile function of the distribution of t_k^* and τ is a user-defined value. (5.37) indicates that τ is the upper limit of α^* .

5.3 Detection Performance with Worldwide Simulations

This section shows the faulty measurement detection results of a set of users distributed over the world during one day. The MATLAB Algorithm Availability Simulation Tool (MAAST) [96] is utilized to simulate code IF combination measurements with tropospheric correction, satellite positions, and user locations. Specifically, we investigate the case of one single fault and multiple faults. For the single fault case, the nominal 24-satellite GPS constellation is used to simulate satellite positions. For

the multiple faults case, the nominal 24-satellite GPS constellation and the nominal 24-satellite Galileo constellation are employed. The sources of almanacs of the two constellations are listed in Table 5.1. The users are placed on a grid every 15 degrees longitude and latitude (which gives 288 locations). For each location, the geometries are simulated every 10 min (which gives 144 time steps). The code IF combination measurements are simulated by adding the randomly generated sample from the given error distribution to the true range. For each time and user location, a given number of measurements are randomly chosen for additional bias injection. Each bias is generated from a uniform distribution in the region $[-20 \text{ m}, -10 \text{ m}] \cup [10 \text{ m}, 20 \text{ m}]$. The SS detector in the baseline ARAIM [20] and the proposed jackknife detector are implemented to detect these faulty measurements, separately. The upper limit τ of the Type I error rate for both detectors is set as 4×10^{-6} . The experiment setting is summarized in Table 5.2.

Table 5.1: Source of almanacs of the GPS and Galileo constellations

Constellation	GPS Week of Almanacs	Source of Almanacs
GPS	2243	U.S. Coast Guard Navigation Center [97]
Galileo	2243	European GNSS Service Center [98]

Table 5.2: Parameters of the fault detection experiments

Number of faults	Constellations	Fault magnitude	τ
1	GPS	$[-20 \text{ m}, -10 \text{ m}] \cup [10 \text{ m}, 20 \text{ m}]$	4×10^{-6}
2	GPS, Galileo	$[-20 \text{ m}, -10 \text{ m}] \cup [10 \text{ m}, 20 \text{ m}]$	4×10^{-6}

The performance of the detectors is evaluated with the actual detection rate at each user location, which is defined as

$$P_{dec} = \frac{\text{Detected epochs in one day}}{\text{Valid epochs in one day}}, \quad (5.39)$$

where the denominator could be less than 144 since the number of satellites in view may not satisfy the minimum requirements for fault detection.

The simulation of the measurement error distribution is detailed in Section 5.3.1. Section 5.3.2 and Section 5.3.3 give the detection results in the single-fault and multiple-fault scenarios, respectively.

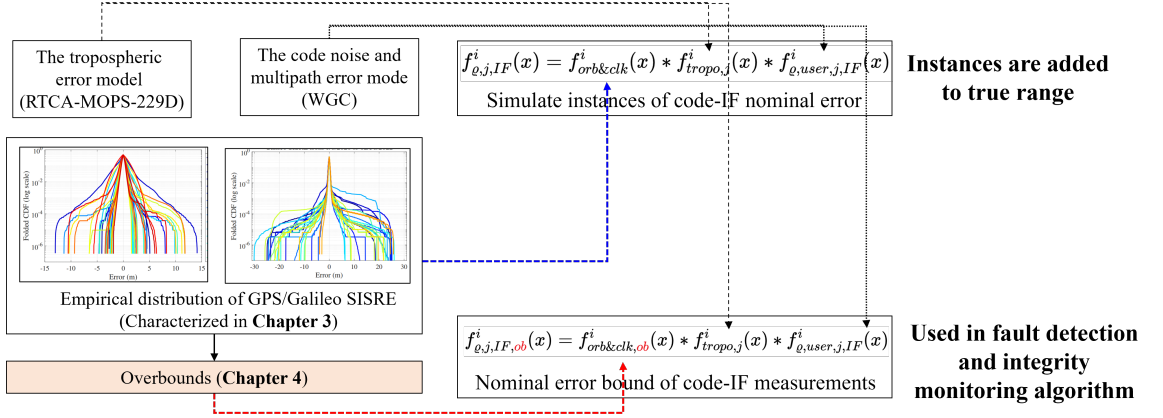


Figure 5.1: The setting of the nominal error and bound of the code IF combination. ‘ob’ represents ‘Gaussian’, ‘acc’, or ‘int’.

5.3.1 Nominal error simulation and bounding

The measurement error of the code IF combination is given in Equation (2.11), which consists the range projection of clock and orbit error, tropospheric error, and multipath and code noise. In this experiment, the distribution of the range projection of clock and orbit error $\varepsilon_{orb\&clk}^i$ is assumed to be the empirical distribution of the $SISRE_{UPE}$ characterized in Chapter 3. The tropospheric error $\varepsilon_{tropo,j}^i$ is assumed to have a zero-mean Gaussian distribution with the 1-sigma bound defined in Equation (2.14). The multipath and code noise $\varepsilon_{\varrho,user,j,IF}^i$ for airborne receivers is assumed to have a zero-mean Gaussian distribution with the 1-sigma bound defined in Appendix A.1. The PDF of the range projection of clock and orbit error, tropospheric error,

and multipath and code noise are denoted as $f_{orb\&clk}^i(x)$, $f_{tropo,j}^i(x)$, and $f_{\rho,user,j,IF}^i(x)$, respectively. For each epoch, the nominal measurement error of the code IF combination is generated by summing up the randomly generated sample from $f_{orb\&clk}^i(x)$, $f_{tropo,j}^i(x)$, and $f_{\rho,user,j,IF}^i(x)$, respectively. Figure 5.1 demonstrates the process of simulating the instance of the nominal measurement error of the code IF combination. Notably, $f_{orb\&clk}^i(x)$ is determined based on authentic experimental data instead of relying on empirical models. This enhances the reliability of the experimental results obtained from simulation.

Two types of nominal error bounds on the code IF combination can be obtained, including the non-Gaussian overbound $f_{\rho,j,IF,acc}^i(x)$ defined in Equation (4.64) and the Gaussian overbound as follows:

$$f_{\rho,j,IF,Gaussian}^i(x) = f_{orb\&clk,Gaussian}^i(x) * f_{tropo,j}^i(x) * f_{\rho,user,j,IF}^i(x), \quad (5.40)$$

where $f_{orb\&clk,Gaussian}^i(x)$ is the Gaussian overbound of the range projection of clock and orbit error with its parameters listed in Tables 3.3 and 3.4. The source of each component in the Gaussian and non-Gaussian overbound is demonstrated in Figure 5.1 .

In the experiment, the Gaussian overbound $f_{\rho,j,IF,Gaussian}^i(x)$ is used for the SS detector. For the jackknife detector, both the Gaussian overbound $f_{\rho,j,IF,Gaussian}^i(x)$ and the non-Gaussian overbound $f_{\rho,j,IF,acc}^i(x)$ are employed. For notations, the jackknife detector using the Gaussian overbound is named as the JKD-Gaussian, while the one using the non-Gaussian overbound is named as the JKD-non-Gaussian. Table 5.3 lists the usage of overbounds in different detectors in the experiment.

Table 5.3: Overbounds used in different detectors

Detector	SS	JKD-Gaussian	JKD-non-Gaussian
Overbounds	$f_{\rho,j,IF,Gaussian}^i(x)$	$f_{\rho,j,IF,Gaussian}^i(x)$	$f_{\rho,j,IF,acc}^i(x)$

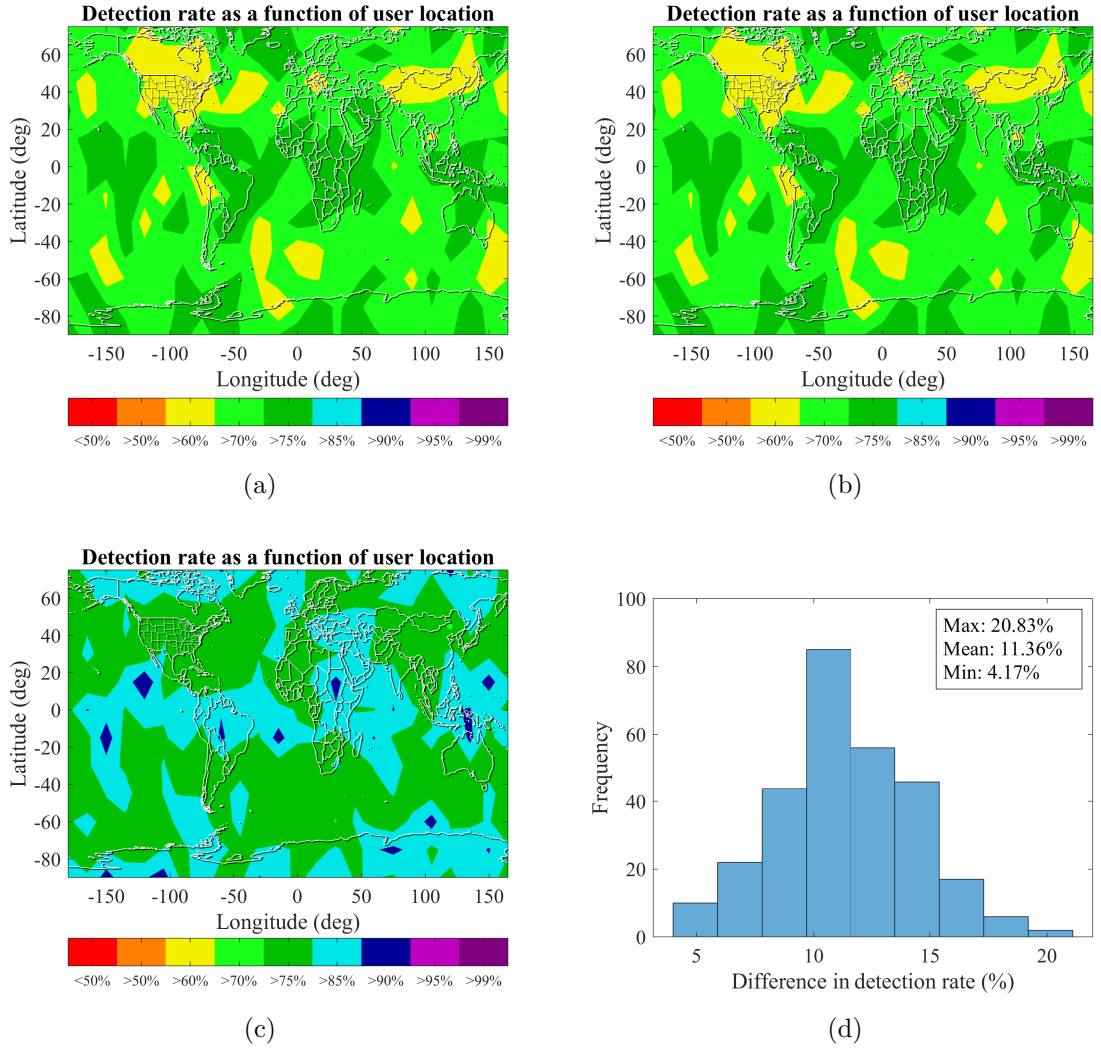


Figure 5.2: Detection rate contour of (a) the SS detector, (b) JKD-Gaussian, and (c) JKD-non-Gaussian with single artificially injected bias. (d) The histogram of the detection rate difference between the JKD-non-Gaussian and the SS detector.

5.3.2 Single-fault detection performance

The first experiment examines the detection performance of the SS detector, JKD-Gaussian, and JKD-non-Gaussian considering a single faulty measurement. Figure 5.2a and Figure 5.2b show the detection rate contour of the SS detector and JKD-Gaussian, respectively. As can be seen, the JKD-Gaussian demonstrates identical

performance as the SS detector, where both detectors show more than 70 % detection rate at most user locations. Figure 5.2c shows the contour plot of the detection rate for the JKD-non-Gaussian. As can be seen, the JKD-non-Gaussian exhibits a substantial enhancement in detection rate when compared to the SS detector. In most user locations, the JKD-non-Gaussian achieves a detection rate of over 75 %. Moreover, in considerable user locations, the detection rate even surpasses 85 %. Figure 5.2d shows the detection rate difference between the JKD-non-Gaussian and the SS detector ($P_{dec}^{\text{JKD-non-Gaussian}} - P_{dec}^{\text{SS}}$) in a histogram view. As can be seen, the maximum improvement in detection rate by the JKD-non-Gaussian exceeds 20 %.

5.3.3 Multiple-fault detection performance

The second experiment examines the detection performance of the SS detector, JKD-Gaussian, and JKD-non-Gaussian considering two faulty measurements. Similar to the single-fault case in Section 5.3.2, the SS detector and JKD-Gaussian show an identical detection performance at all user locations, as shown in Figures 5.3a and 5.3b. However, the detection rate of these two methods is less than 70 % at most user locations, which is significantly smaller than that in the single-fault case (see Figures 5.2a and 5.2b). The primary reason is that the two-fault setting uses the dual constellation (GPS and Galileo) to simulate code IF combination measurements, where the Galileo SISRE is featured with significant heavy-tailed properties. As discussed in Section 4.4, this heavy-tailed property excessively enlarges the standard deviation of the Gaussian overbounds of SISRE. Such over-conservative Gaussian overbounds can inflate the type II error of hypothesis testing, which corresponds to the degradation of the detection rate in the dual-constellation setting. However, the JKD-non-Gaussian still shows satisfactory detection performance in the dual-constellation setting, where the detection rate is larger than 85 % in most user locations, as shown in Figure 5.3c. This result again emphasizes the benefits of using PGO for heavy-tailed error bounding in detection tasks. Figure 5.3d shows the detection rate difference

between the JKD-non-Gaussian and the SS detector in the two-fault setting. As can be seen, the maximum improvement in detection rate by the JKD-non-Gaussian exceeds 34 %.

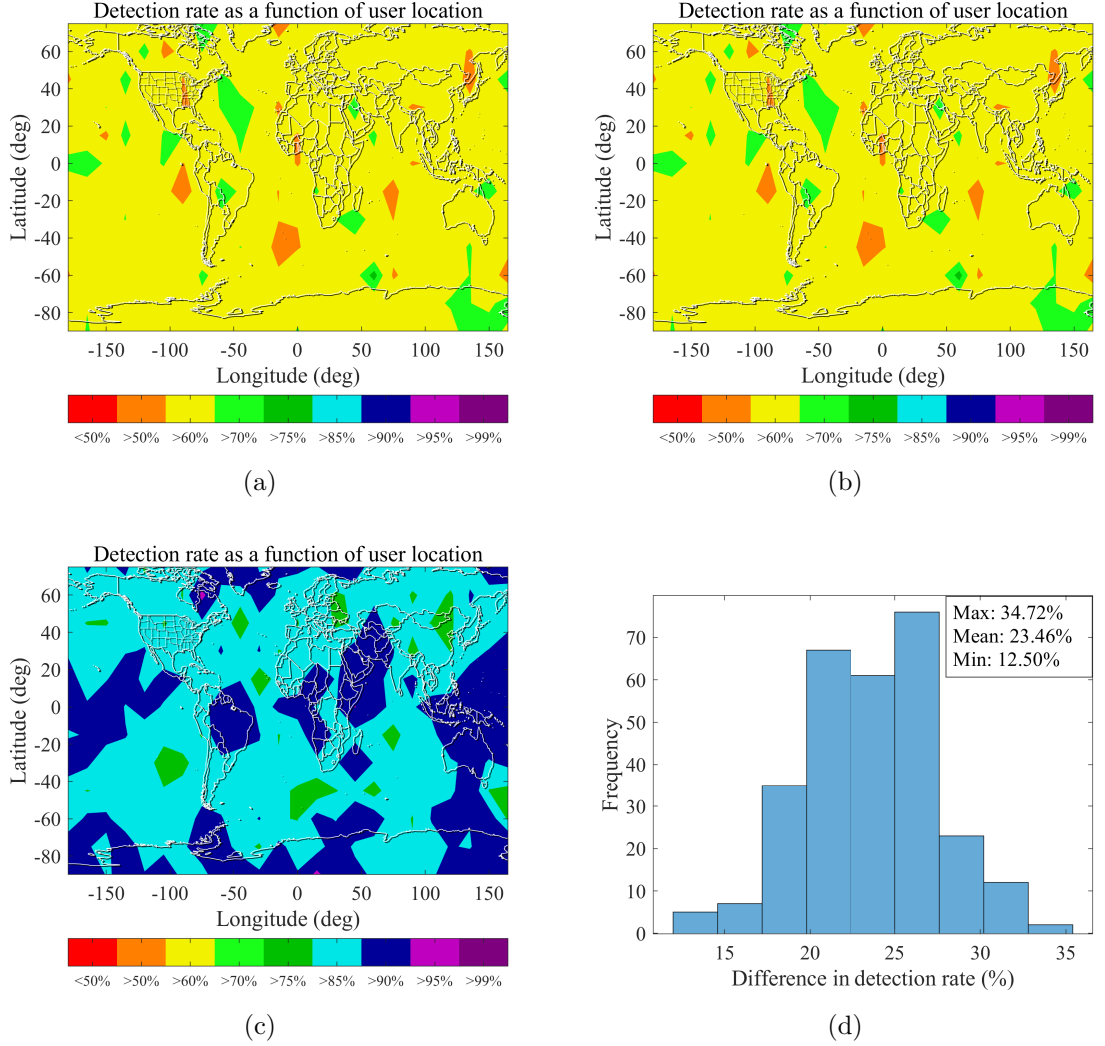


Figure 5.3: Detection rate contour of (a) the SS detector, (b) JKD-Gaussian, and (c) JKD-non-Gaussian with two artificially injected biases. (d) The histogram of the detection rate difference between the JKD-non-Gaussian and the SS detector.

The single-fault and multi-fault detection experiments demonstrate the superiority of the JKD-non-Gaussian. Shortly speaking, this superiority can be primarily attributed to two factors: 1) a sharper overbound for heavy-tailed measurement error and 2) the

accommodation of the jackknife test for non-Gaussian overbounds. On the one hand, the sharper overbound provided by PGO can better characterize the measurement error distribution than the Gaussian overbound, as shown in Section 4.4, providing an accurate probabilistic model for hypothesis testing. It is known that an inaccurate probabilistic model can hinder the performance of hypothesis testing, such as inflating the type I error (false alarm) or type II error (miss detection), and even make the testing invalid. On the other hand, the theoretical underpinnings of the jackknife test provide a solid basis for its application in fault detection with non-Gaussian overbounds. The threshold for the JKD-non-Gaussian is derived through rigorous mathematical derivations rather than relying on simulation techniques. These two factors complement each other, resulting in the exceptional performance of the JKD-non-Gaussian. In the next chapter, the JKD-non-Gaussian is further exploited to develop an integrity monitoring algorithm to support integrity applications under stringent navigation requirements.

Chapter 6

Integrity Monitoring with Non-Gaussian Nominal Error Bound

The non-Gaussian nominal error bound constructed in Chapter 4 and the non-Gaussian fault detection methods proposed in Chapter 5 lay the foundation for this chapter to develop a multiple-hypothesis-based integrity monitoring algorithm with non-Gaussian nominal error bounds. The proposed method is named the jackknife ARAIM algorithm to emphasize its usage of the jackknife detector. The jackknife ARAIM algorithm follows a similar process to the baseline ARAIM algorithm, beginning with defining the threat model, constructing the fault detectors, and determining their threshold to comply with the continuity requirements, then evaluating integrity risks, and concluding with deriving protection levels. The principal difference between the proposed jackknife ARAIM algorithm and the baseline ARAIM algorithm lies in the choice of fault detectors. Instead of using solution separation in the position domain, the proposed method systematically exploits the properties of the jackknife detector in the range domain and derives a tight bound of the integrity risk. The

proposed method extends the scope of the integrity monitoring algorithm to handle non-Gaussian nominal error bounds.

6.1 Threat Model and Multiple Hypotheses

The proposed method uses the same threat model defined in the baseline ARAIM algorithm [20], as elaborated in Section 2.3.3. Briefly, the threat model defines a collection of error modes that partition the whole measurement space [18, 99]. The fault mode 0 represents the fault-free case while other fault modes (i.e., $1, 2, \dots, N_{\text{fault modes}}$) indicate the presence of single or multiple faults. The total number of fault modes, i.e., $N_{\text{fault modes}} + 1$, is determined by the maximum number of simultaneous faults k_{max} . Based on the threat model, the following hypotheses are constructed:

- H_0 : The fault-free hypothesis (fault mode 0).
- H_k : The hypothesis corresponding to fault mode k .

This set of hypotheses is the same as that in the jackknife detector developed in Section 5.2. Therefore, the jackknife detector can be directly used as the monitor in the integrity monitoring algorithm.

6.2 Determine the Threshold of Monitors

The threshold of monitors, i.e., jackknife detectors, is determined so that the continuity requirement is satisfied. The continuity risk in Equation (2.34) can be rewritten as follows:

$$P_{\text{continuity}} = P\left(\bigcup_{k=1}^{N_{\text{fault modes}}} t_k^* \in R_k \mid H_0\right) P_{H_0} \leq C_{\text{REQ,FA}}, \quad (6.1)$$

with R_k given by

$$R_k = \{t_k^* \mid |t_k^*| \geq T_k\}, k = 1, 2, \dots, N_{\text{fault modes}}. \quad (6.2)$$

Since $R_1, R_2, \dots, R_{N_{\text{fault modes}}}$ are not mutually exclusive, we have

$$P_{\text{continuity}} \leq \sum_{k=1}^{N_{\text{fault modes}}} P(t_k^* \in R_k \mid H_0) P_{H_0} = \sum_{k=1}^{N_{\text{fault modes}}} P(|t_k^*| \geq T_k \mid H_0) P_{H_0}. \quad (6.3)$$

The threshold T_k is determined by the allocated continuity budget caused by false alert

$$T_k = Q_{t_k^*}^{-1} \left(\frac{C_{\text{REQ,FA}}}{2N_{\text{fault modes}} P_{H_0}} \right). \quad (6.4)$$

As shown in Equations (5.20), (5.30), and (5.33), t_k^* is the linear combination of nominal measurement error bounds, i.e., $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$. Here, $\varepsilon_j, j = 1, 2, \dots, n$ refers to the nominal error bound for accuracy, which is defined in Equation (4.64). The quantile function $Q_{t_k^*}^{-1}(\cdot)$ can be evaluated by using the numerical method developed in Section 4.3.5.3.

In Equation (6.4), the equal allocation strategy of the continuity budget is adopted, which is the same as that in the baseline ARAIM algorithm. However, Equation (6.4) does not require the partition of vertical and horizontal components of the continuity budget, which is done in the baseline ARAIM algorithm in Equation (2.49).

6.3 Integrity Risk Evaluation

The detection threshold determined in Equation (6.4) can be used to evaluate the integrity risk in Equation (2.33). By using the jackknife detector developed in Chapter 5, Equation (2.33) can be written by

$$P_{\text{HMI}} = \sum_{i=0}^{N_{\text{fault modes}}} P \left(\{|e_0| > \ell\} \cap \bigcap_{k=1}^{N_{\text{fault modes}}} |t_k^*| < T_k \mid H_k \right) P_{H_i} + P_{\text{not monitored}} \leq I_{\text{REQ}}. \quad (6.5)$$

In the navigation system, the estimation error on the parameter of interest is the positioning error. Therefore, e_0 can be rewritten as

$$e_0 = (\hat{\mathbf{x}} - \mathbf{x})_v, \quad (6.6)$$

where the subscript $v = 1, 2, 3$ designates the east, north, and up components of the position error, respectively. The alert limit corresponding to $(\hat{\mathbf{x}} - \mathbf{x})_v$ is denoted as ℓ_v in the following derivations.

Let

$$I_{cal} = \sum_{i=0}^{N_{\text{fault modes}}} P \left(\{ |(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \} \cap \bigcap_{k=1}^{N_{\text{fault modes}}} |t_k^*| < T_k \mid H_i \right) P_{H_i}, \quad (6.7)$$

which is the sum of hazardously misleading information (HMI) probabilities over the fault-free hypothesis and other faulted hypotheses.

6.3.1 Bound on the probability of HMI under H_0

In the fault-free hypothesis H_0 , a bound on the probability of HMI is established as follows

$$P \left(\{ |(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \} \cap \bigcap_{k=1}^{N_{\text{fault modes}}} |t_k^*| < T_k \mid H_0 \right) \leq P (|(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \mid H_0). \quad (6.8)$$

This bound is obtained by ignoring knowledge of no detection, which can be considered a tight bound [18]. This is because the probability of no detection under the fault-free hypothesis is larger than $1 - C_{\text{REQ,FA}}$, as ensured by Equation (6.1).

By substituting Equation (5.1) and Equation (5.2) into $(\hat{\mathbf{x}} - \mathbf{x})_v$, we have

$$(\hat{\mathbf{x}} - \mathbf{x})_v = (\mathbf{S}\boldsymbol{\varepsilon})_v = \sum_{i=1}^n S_{v,i} \varepsilon_i, \quad (6.9)$$

where $S_{v,i}$ is the (v, i) th element in $\mathbf{S}$. Then the PDF of $(\hat{\mathbf{x}} - \mathbf{x})_v$ is given by (see Appendix A.6)

$$f_{(\hat{\mathbf{x}} - \mathbf{x})_v}(t) = \prod_{j=1}^n |S_{v,j}|^{-1} f_{\varepsilon_1} \left(\frac{t}{|S_{v,1}|} \right) * f_{\varepsilon_2} \left(\frac{t}{|S_{v,2}|} \right) * \dots * f_{\varepsilon_n} \left(\frac{t}{|S_{v,n}|} \right). \quad (6.10)$$

Equation (6.10) can be used to evaluate the bound in Equation (6.8).

6.3.2 Bound on the probability of HMI under H_k

In the faulted hypothesis H_k , a similar bound on the probability of HMI is given as follows:

$$\begin{aligned} & P \left(\{ |(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \} \cap \bigcap_{k=1}^{N_{\text{fault modes}}} |t_k^*| < T_k \mid H_k \right) \\ & \leq P \left(\{ |(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \} \cap \{ |t_k^*| < T_k \} \mid H_k \right) . \end{aligned} \quad (6.11)$$

Again, this bound is obtained by ignoring knowledge of no detection for all other hypothesis tests, except for the one for the test H_0 v.s. H_k . As proven in [18], Equation (6.11) also provides a tight bound on the probability of HMI under H_k . The right-hand-side of Equation (6.11) can be simplified by invoking the conditional probability

$$\begin{aligned} & P \left(\{ |(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \} \cap \{ |t_k^*| < T_k \} \mid H_k \right) \\ & = P \left(|(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \mid H_k \cap \{ |t_k^*| < T_k \} \right) P \left(|t_k^*| < T_k \mid H_k \right) \\ & \leq P \left(|(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \mid H_k \cap \{ |t_k^*| < T_k \} \right) . \end{aligned} \quad (6.12)$$

The inequality in the second line bounds $P(|t_k^*| < T_k \mid H_k)$ with $P(|t_k^*| < T_k \mid H_k) = 1$.

A further relaxation of Equation (6.12) is achieved by exploiting the structure of $(\hat{\mathbf{x}} - \mathbf{x})_v$ under H_k . Define the fault vector in the faulted hypothesis H_k as $\mathbf{b}^{(k)}$. This $n \times 1$ vector takes the following form:

$$b_j^{(k)} = \begin{cases} b_j & \text{if } j \in \text{idx}_k^{ex} \\ 0 & \text{otherwise} \end{cases} , \quad (6.13)$$

where $b_j^{(k)}$ is the j th element of $\mathbf{b}^{(k)}$ and $b_j, j = 1, 2, \dots, n$ is an unknown constant with non-zero values. In the faulted hypothesis H_k , the linearized measurement model in Equation (2.20) can be written by

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \boldsymbol{\varepsilon} + \mathbf{b}^{(k)} , \quad (6.14)$$

where

$$y_j = \begin{cases} \mathbf{g}_j \mathbf{x} + \varepsilon_j + b_j & \text{if } j \in idx_k^{ex} \\ \mathbf{g}_j \mathbf{x} + \varepsilon_j & \text{otherwise,} \end{cases} \quad (6.15)$$

and $\mathbf{g}_j$ is the j th row of $\mathbf{G}$.

Different from Section 6.2, $\varepsilon_i, i = 1, 2, \dots, n$ in Equations (6.14) and (6.15) refers to the nominal measurement error bound for integrity, which is developed in Section 4.65. This kind of bound considers the effects of nominal signal deformation errors, which is realized by introducing a b_{nom} term to create two equally shifted nominal measurement error bounds for accuracy. To simplify the derivation, we first ignore the effects of nominal signal deformation errors by setting $b_{nom,i} = 0, i = 1, 2, \dots, n$. Then the nominal measurement error bound for integrity is the same as that for accuracy.

Now, $(\hat{\mathbf{x}} - \mathbf{x})_v$ under H_k can be written by

$$\begin{aligned} (\hat{\mathbf{x}} - \mathbf{x})_v \mid H_k &= (\mathbf{S}\mathbf{y} - \mathbf{x})_v \mid H_k \\ &= (\mathbf{S}(\mathbf{G}\mathbf{x} + \boldsymbol{\varepsilon} + \mathbf{b}^{(k)}) - \mathbf{x})_v \\ &= (\mathbf{S}\boldsymbol{\varepsilon} + \mathbf{S}\mathbf{b}^{(k)})_v \\ &= \sum_{i=1}^n S_{v,i} \varepsilon_i + \sum_{j \in idx_k^{ex}} S_{v,j} b_j. \end{aligned} \quad (6.16)$$

For each $j \in idx_k^{ex}$, the corresponding jackknife residual is given by

$$\begin{aligned} t_j^{(k)} &= y_j - \hat{y}_j \\ &= \mathbf{g}_j \mathbf{x} + \varepsilon_j + b_j - \mathbf{g}_j \hat{\mathbf{x}}^{(k)} \\ &= \mathbf{g}_j (\mathbf{x} - \hat{\mathbf{x}}^{(k)}) + \varepsilon_j + b_j \\ &= -\mathbf{g}_j \mathbf{S}^{(k)} \boldsymbol{\varepsilon} + \varepsilon_j + b_j. \end{aligned} \quad (6.17)$$

The last line holds because $\hat{\mathbf{x}}^{(k)} - \mathbf{x} = \mathbf{S}^{(k)} \boldsymbol{\varepsilon}$. Then, we have

$$b_j = t_j^{(k)} + \mathbf{g}_j \mathbf{S}^{(k)} \boldsymbol{\varepsilon} - \varepsilon_j. \quad (6.18)$$

By substituting Equation (6.18) into Equation (6.16), we have

$$\begin{aligned}
 (\hat{\mathbf{x}} - \mathbf{x})_v \mid H_k &= \sum_{i=1}^n S_{v,i} \varepsilon_i + \sum_{j \in id\mathbf{x}_k^{ex}} S_{v,j} (t_j^{(k)} + \mathbf{g}_j \mathbf{S}^{(k)} \boldsymbol{\varepsilon} - \varepsilon_j) \\
 &= \sum_{j \notin id\mathbf{x}_k^{ex}} S_{v,j} \varepsilon_j + \sum_{j \in id\mathbf{x}_k^{ex}} S_{v,j} \mathbf{g}_j \mathbf{S}^{(k)} \boldsymbol{\varepsilon} + \sum_{j \in id\mathbf{x}_k^{ex}} S_{v,j} t_j^{(k)}.
 \end{aligned} \tag{6.19}$$

Let $\mathbf{E}^{(k)}$ be a $n \times n$ diagonal matrix with the following definition

$$E_{j,j}^{(k)} = \begin{cases} 0 & \text{if } j \in id\mathbf{x}_k^{ex} \\ 1 & \text{otherwise} \end{cases}. \tag{6.20}$$

Equation (6.19) can be simplified to

$$(\hat{\mathbf{x}} - \mathbf{x})_v \mid H_k = \mathbf{q}^{(k)} \boldsymbol{\varepsilon} + \sum_{j \in id\mathbf{x}_k^{ex}} S_{v,j} t_j^{(k)}, \tag{6.21}$$

where

$$\mathbf{q}^{(k)} = \mathbf{s}_v \mathbf{E}^{(k)} + \sum_{j \in id\mathbf{x}_k^{ex}} S_{v,j} \mathbf{g}_j \mathbf{S}^{(k)}. \tag{6.22}$$

The distribution of $\mathbf{q}^{(k)} \boldsymbol{\varepsilon}$ is given by (see Appendix A.6)

$$f_{\mathbf{q}^{(k)} \boldsymbol{\varepsilon}}(x) = \prod_{j=1}^n |q_j^{(k)}|^{-1} f_{\varepsilon_1} \left(\frac{x}{|q_1^{(k)}|} \right) * f_{\varepsilon_2} \left(\frac{x}{|q_2^{(k)}|} \right) * \dots * f_{\varepsilon_n} \left(\frac{x}{|q_n^{(k)}|} \right), \tag{6.23}$$

where $q_j^{(k)}, j = 1, 2, \dots, n$ is the j th element of $\mathbf{q}^{(k)}$.

Then the bound on the probability of HMI under H_k in Equation (6.12) can be written by

$$P(|(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \mid H_k \cap \{|t_k^*| < T_k\}) \tag{6.24a}$$

$$= P \left(|\mathbf{q}^{(k)} \boldsymbol{\varepsilon} + \sum_{j \in id\mathbf{x}_k^{ex}} S_{v,j} t_j^{(k)}| > \ell_v \mid H_k \cap \{|t_k^*| < T_k\} \right) \tag{6.24b}$$

$$\leq P \left(|\mathbf{q}^{(k)} \boldsymbol{\varepsilon}| + \left| \sum_{j \in id\mathbf{x}_k^{ex}} S_{v,j} t_j^{(k)} \right| > \ell_v \mid H_k \cap \{|t_k^*| < T_k\} \right) \tag{6.24c}$$

The second line holds because of the triangular inequality.

When $k \leq n$, $t_k^* = t_k$. Then, the right-hand-side of Equation (6.24) can be written by

$$P(|\mathbf{q}^{(k)}\boldsymbol{\varepsilon}| + |S_{v,k}t_k| > \ell_v \mid H_k \cap \{t_k \leq T_k\}) \leq P(|\mathbf{q}^{(k)}\boldsymbol{\varepsilon}| + |S_{v,k}|T_k > \ell_v \mid H_k) . \quad (6.25)$$

When $k > n$, $t_k^* = \tilde{t}_k = \sum_{j \in id x_k^{ex}} S_{v,j} t_j^{(k)}$. Then, the right-hand-side of Equation (6.24) can be written by

$$\begin{aligned} & P \left(|\mathbf{q}^{(k)}\boldsymbol{\varepsilon}| + \left| \sum_{j \in id x_k^{ex}} S_{v,j} t_j^{(k)} \right| > \ell_v \mid H_k \cap \left\{ \left| \sum_{j \in id x_k^{ex}} S_{v,j} t_j^{(k)} \right| < T_k \right\} \right) \\ & \leq P(|\mathbf{q}^{(k)}\boldsymbol{\varepsilon}| + T_k > \ell_v \mid H_k) . \end{aligned} \quad (6.26)$$

6.3.3 Finalized bound of integrity risk

Finally, the bound of integrity risk for monitored fault modes in Equation (6.7) is given by summarizing Equations (6.8), (6.25) and (6.26) as follows:

$$\begin{aligned} I_{cal} & \leq P(|(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \mid H_0) P_{H_0} \\ & \quad + \sum_{k=1}^n P(|\mathbf{q}^{(k)}\boldsymbol{\varepsilon}| + |S_{v,k}|T_k > \ell_v \mid H_k) P_{H_k} \\ & \quad + \sum_{k=n+1}^{N_{\text{fault modes}}} P(|\mathbf{q}^{(k)}\boldsymbol{\varepsilon}| + T_k > \ell_v \mid H_k) P_{H_k} \\ & \leq I_{\text{REQ}}^v \left(1 - \frac{P_{\text{not monitored}}}{I_{\text{REQ}}} \right) , \end{aligned} \quad (6.27)$$

where I_{REQ}^v , $v = 1, 2, 3$ has the same definition as in Section 2.3.5 and stands for the integrity budget for different components. Notably, the distributions of $(\hat{\mathbf{x}} - \mathbf{x})_v$ and $\mathbf{q}^{(k)}\boldsymbol{\varepsilon}$ are known, as given in Equation (6.10) and Equation (6.23), respectively. Hence, the inequality condition in the last line can be evaluated to check if the integrity requirement is satisfied.

So far, we have derived the bound of integrity risk for monitored fault modes with $b_{nom} = 0$. To consider the effects of nominal signal deformation errors, the integrity

risk bound in Equation (6.27) can be modified as follows:

$$\begin{aligned}
 I_{cal} &\leq P \left(|(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v - b_v^{(0)} \mid H_0 \right) P_{H_0} \\
 &+ \sum_{k=1}^n P \left(|\mathbf{q}^{(k)} \boldsymbol{\varepsilon}| + |S_{v,k}| T_k > \ell_v - b_v^{(k)} \mid H_k \right) P_{H_k} \\
 &+ \sum_{k=n+1}^{N_{\text{fault modes}}} P \left(|\mathbf{q}^{(k)} \boldsymbol{\varepsilon}| + T_k > \ell_v - b_v^{(k)} \mid H_k \right) P_{H_k} \\
 &\leq I_{\text{REQ}}^v \left(1 - \frac{P_{\text{not monitored}}}{I_{\text{REQ}}} \right),
 \end{aligned} \tag{6.28}$$

where $b_v^{(k)}$ represents the worst-case impact of nominal signal deformation errors on the position solution:

$$b_v^{(k)} = \sum_{i=1}^n |S_{v,i}^{(k)}| b_{\text{nom},i}. \tag{6.29}$$

6.4 Protection Level Derivation

As an alternative to integrity risk evaluation, PL can be derived from Equation (6.28) by replacing the alert limit ℓ_v with protection level PL_v and replacing the last inequality with equality as follows:

$$\begin{aligned}
 &P \left(|(\hat{\mathbf{x}} - \mathbf{x})_v| > PL_v - b_v^{(0)} \mid H_0 \right) P_{H_0} \\
 &+ \sum_{k=1}^n P \left(|\mathbf{q}^{(k)} \boldsymbol{\varepsilon}| + |S_{v,k}| T_k > PL_v - b_v^{(k)} \mid H_k \right) P_{H_k} \\
 &+ \sum_{k=n+1}^{N_{\text{fault modes}}} P \left(|\mathbf{q}^{(k)} \boldsymbol{\varepsilon}| + T_k > PL_v - b_v^{(k)} \mid H_k \right) P_{H_k} \\
 &= I_{\text{REQ}}^v \left(1 - \frac{P_{\text{not monitored}}}{I_{\text{REQ}}} \right).
 \end{aligned} \tag{6.30}$$

To solve PL_v , the integrity budget $I_{\text{REQ}}^v \left(1 - \frac{P_{\text{not monitored}}}{I_{\text{REQ}}}\right)$ needs to be allocated to each fault mode. Specifically, PL_v is given by

$$PL_v = \max \left\{ Q_{(\hat{\mathbf{x}} - \mathbf{x})_v}^{-1} \left(\frac{I_{REQ,0}^v}{2P_{H_0}} \right) + b_v^{(0)}, \max_{1 \leq k \leq n} \left\{ Q_{\mathbf{q}^{(k)}\epsilon}^{-1} \left(\frac{I_{REQ,k}^v}{2P_{H_k}} \right) + |S_{v,k}|T_k + b_v^{(k)} \right\}, \right. \\ \left. \max_{n < k \leq N_{\text{fault modes}}} \left\{ Q_{\mathbf{q}^{(k)}\epsilon}^{-1} \left(\frac{I_{REQ,k}^v}{2P_{H_k}} \right) + T_k + b_v^{(k)} \right\} \right\}, \quad (6.31)$$

where

$$\sum_{k=1}^{N_{\text{fault modes}}} I_{REQ,k}^v = I_{\text{REQ}}^v \left(1 - \frac{P_{\text{not monitored}}}{I_{\text{REQ}}}\right). \quad (6.32)$$

The quantile functions $Q_{(\hat{\mathbf{x}} - \mathbf{x})_v}^{-1}$ and $Q_{\mathbf{q}^{(k)}\epsilon}^{-1}$ can be evaluated by using the numerical method developed in Section 4.3.5.3.

In this thesis, the equal allocation strategy for integrity is applied as follows:

$$I_{REQ,k}^v = \frac{1}{N_{\text{fault modes}}} I_{\text{REQ}}^v \left(1 - \frac{P_{\text{not monitored}}}{I_{\text{REQ}}}\right). \quad (6.33)$$

The VPL is directly given by PL_3 , i.e.,

$$VPL = PL_3, \quad (6.34)$$

and the HPL is given by synthesizing PL_1 and PL_2 as follows:

$$HPL = \sqrt{PL_1^2 + PL_2^2}. \quad (6.35)$$

6.5 Consideration of Constellation Faults

As discussed in Section 5.2.1, the jackknife residual is not computable in the constellation fault mode. Therefore, the PL calculation in Section 6.4 does not consider constellation fault modes. However, it is essential to consider the possibility of constellation faults in the multi-constellation system to protect integrity. To address this issue, one can use the solution separation detector to construct the hypothesis regarding the constellation fault and integrate it into the PL equations in Section 6.4.

Let Ω_{const} be the set of fault modes involving constellation faults. Under each fault mode $k \in \Omega_{\text{const}}$, the integrity risk of HMI is given by

$$P \left(\{ |(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \} \cap \{ |d_v^{(k)}| < D_{k,v} \} \mid H_k, k \in \Omega_{\text{const}} \right), \quad (6.36)$$

where $d_v^{(k)}$ and $D_{k,v}$ are the solution separation test statistic and its threshold, respectively (see Section 2.3.4). According to the triangular inequality,

$$|(\hat{\mathbf{x}} - \mathbf{x})_v| = |(\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(k)} + \hat{\mathbf{x}}^{(k)} - \mathbf{x})_v| \leq |(\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(k)})_v| + |(\hat{\mathbf{x}}^{(k)} - \mathbf{x})_v|. \quad (6.37)$$

Therefore, Equation (6.36) can be bounded by

$$P \left(\{ |(\hat{\mathbf{x}} - \mathbf{x})_v| > \ell_v \} \cap \{ |d_v^{(k)}| < D_{k,v} \} \mid H_k, k \in \Omega_{\text{const}} \right) \quad (6.38a)$$

$$\leq P \left(\{ |(\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(k)})_v| + |(\hat{\mathbf{x}}^{(k)} - \mathbf{x})_v| > \ell_v \} \cap \{ |d_v^{(k)}| < D_{k,v} \} \mid H_k, k \in \Omega_{\text{const}} \right) \quad (6.38b)$$

$$\leq P \left(|(\hat{\mathbf{x}}^{(k)} - \mathbf{x})_v| + D_{k,v} > \ell_v \mid H_k, k \in \Omega_{\text{const}} \right). \quad (6.38c)$$

Following the steps in Section 6.3.3, Equation (6.30) can be eventually re-written as

$$\begin{aligned} & P \left(|(\hat{\mathbf{x}} - \mathbf{x})_v| > PL_v - b_v^{(0)} \mid H_0 \right) P_{H_0} \\ & + \sum_{k=1}^n P \left(|\mathbf{q}^{(k)} \boldsymbol{\varepsilon}| + |S_{v,k}| T_k > PL_v - b_v^{(k)} \mid H_k \right) P_{H_k} \\ & + \sum_{k=n+1, k \notin \Omega_{\text{const}}}^{N_{\text{fault modes}}} P \left(|\mathbf{q}^{(k)} \boldsymbol{\varepsilon}| + T_k > PL_v - b_v^{(k)} \mid H_k \right) P_{H_k} \\ & + \sum_{k \in \Omega_{\text{const}}} P \left(|(\hat{\mathbf{x}}^{(k)} - \mathbf{x})_v| + D_{k,v} > PL_v - b_v^{(k)} \mid H_k \right) P_{H_k} \\ & = I_{\text{REQ}}^v \left(1 - \frac{P_{\text{not monitored}}}{I_{\text{REQ}}} \right). \end{aligned} \quad (6.39)$$

Notably, the last term in the left-hand-side of Equation (6.39) is obtained using the solution separation scheme, which assumes that the nominal error is Gaussian bounded. Therefore, the distribution of $(\hat{\mathbf{x}}^{(k)} - \mathbf{x})_v \mid k \in \Omega_{\text{const}}$ is given by

$$(\hat{\mathbf{x}}^{(k)} - \mathbf{x})_v \mid k \in \Omega_{\text{const}} \sim \mathcal{N} \left(0, (\sigma_v^{(k)})^2 \right), \quad (6.40)$$

where $\sigma_v^{(k)}$ is the standard deviation of the k th subset solution. Similarly, $D_{k,v}$ is also determined with the Gaussian nominal error bound, as detailed in Equation (2.49).

Finally, with the equal allocation strategy on the integrity budget, the PL can be obtained by

$$\begin{aligned}
 PL_v = \max & \left\{ Q_{(\hat{\mathbf{x}}-\mathbf{x})_v}^{-1} \left(\frac{I_{REQ,0}^v}{2P_{H_0}} \right) + b_v^{(0)}, \max_{1 \leq k \leq n} \left\{ Q_{\mathbf{q}^{(k)}\boldsymbol{\varepsilon}}^{-1} \left(\frac{I_{REQ,k}^v}{2P_{H_k}} \right) + |S_{v,k}|T_k + b_v^{(k)} \right\}, \right. \\
 & \max_{n < k \leq N_{\text{fault modes}, k \notin \Omega_{\text{const}}}} \left\{ Q_{\mathbf{q}^{(k)}\boldsymbol{\varepsilon}}^{-1} \left(\frac{I_{REQ,k}^v}{2P_{H_k}} \right) + T_k + b_v^{(k)} \right\}, \\
 & \left. \max_{k \in \Omega_{\text{const}}} \left\{ \sigma_v^{(k)} Q^{-1} \left(\frac{I_{REQ,k}^v}{2P_{H_k}} \right) + D_{k,v} + b_v^{(k)} \right\} \right\}.
 \end{aligned} \tag{6.41}$$

To sum up, Algorithm 3 lists the steps for implementing the proposed jackknife ARAIM.

6.6 Worldwide Simulation

This section conducts a worldwide simulation to evaluate the performance of the proposed jackknife ARAIM algorithm. Specifically, the MAAST toolset [96] is used to simulate code IF combination measurements with tropospheric correction, satellite positions, and user locations. The simulation has the same setting as that in the fault detection experiments in Section 5.3, producing 288×144 location-time geometries in one day. Both the single constellation (the nominal 24-satellite GPS constellation) and dual constellations (the aforementioned GPS constellations and the nominal 24-satellite Galileo constellation) cases are examined, where the almanacs file is defined in Table 5.1. The nominal measurement error of the code IF combination is generated by summing up the randomly generated sample from $f_{orb\&clk}^i(x)$, $f_{tropo,j}^i(x)$, and $f_{\rho,user,j,IF}^i(x)$. As discussed in Section 5.3.1, such the simulation setting can enhance the reliability of experimental results.

The proposed jackknife ARAIM algorithm is compared with the baseline ARAIM algorithm [16]. For the baseline ARAIM algorithm, the Gaussian overbound $f_{\rho,j,IF,Gaussian}^i(x)$ is used. For the jackknife ARAIM algorithm, both the Gaussian overbound $f_{\rho,j,IF,Gaussian}^i(x)$

Algorithm 3 Implementation of Jackknife ARAIM**Input:**

Measurements and geometry matrix: $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$, $\mathbf{G}$
 Number of constellations: N_{const}
 ISM parameters: $P_{\text{sat}}, P_{\text{const}}, P_{\text{THRES}}$
 Continuity and Integrity budgets: $C_{\text{REQ,FA}}, I_{\text{REQ}}^v$
 Overbound of nominal code IF combination error: $f_{\varepsilon_j}, \forall j = 1, 2, \dots, n$
 Nominal signal deformation errors: b_{nom}

Output:

Detection state, VPL, HPL

- 1: $k_{\text{max}}, P_{\text{not monitored}} \leftarrow$ Equations (A.10) and (A.12) with $P_{\text{sat}}, P_{\text{const}}, P_{\text{THRES}}, n$ and N_{const}
- 2: $N_{\text{fault modes}} \leftarrow$ Equation (2.32) with k_{max} and n
- 3: $P_{H_i}, \forall i = 1, 2, \dots, N_{\text{fault modes}} \leftarrow$ Equation (A.5) with P_{sat} and P_{const}
- 4: $\hat{\mathbf{x}}_v, \mathbf{S} \leftarrow$ Equations (2.36) and (2.35) with $\mathbf{y}$ and $\mathbf{G}$
- 5: **for** $k = 1 : N_{\text{fault modes}}$ **do**
- 6: $\hat{\mathbf{x}}_v^{(k)}, \mathbf{S}^{(k)} \leftarrow$ Equations (2.39) and (2.38) with $\mathbf{y}$ and $\mathbf{G}$
- 7: $\sigma_v^{(k)} \leftarrow$ Equation (2.41) with $\mathbf{G}$ and $f_{\varepsilon_j}, \forall j = 1, 2, \dots, n$
- 8: **if** $k \notin \Omega_{\text{const}}$ **then**
- 9: $t_k^* \leftarrow$ Equation (5.33) with $\mathbf{y}, \hat{\mathbf{x}}^{(k)}, \mathbf{S}$, and $\mathbf{S}^{(k)}$
- 10: $T_k \leftarrow$ Equation (6.4) with $C_{\text{REQ,FA}}, N_{\text{fault modes}}, P_{H_0}$, and $f_{\varepsilon_j}, \forall j = 1, 2, \dots, n$
- 11: **else**
- 12: $d_v^{(k)} \leftarrow$ Equation (2.43) with $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}^{(k)}$
- 13: $\sigma_{ss,v}^{(k)} \leftarrow$ Equation (2.45) with $\mathbf{S}, \mathbf{S}^{(k)}$, and $f_{\varepsilon_j}, \forall j = 1, 2, \dots, n$
- 14: $D_{k,v} \leftarrow$ Equation (2.49) with $\sigma_{ss,v}^{(k)}, C_{\text{REQ,FA}}, N_{\text{fault modes}}$, and P_{H_0}
- 15: **end if**
- 16: **if** $t_k^* > T_k$ or $d_v^{(k)} > D_{k,v}$ **then**
- 17: Detect faults and alert users!
- 18: Return
- 19: **end if**
- 20: **end for**
- 21: $I_{\text{REQ},0}^v \leftarrow$ Equation (6.33) with $I_{\text{REQ}}^v, N_{\text{fault modes}}$ and $P_{\text{not monitored}}$
- 22: $b_v^{(0)} \leftarrow$ Equation (6.29) with $\mathbf{S}$ and b_{nom}
- 23: **for** $k = 1 : N_{\text{fault modes}}$ **do**
- 24: $I_{\text{REQ},k}^v \leftarrow$ Equation (6.33) with $I_{\text{REQ}}^v, N_{\text{fault modes}}$ and $P_{\text{not monitored}}$
- 25: $\mathbf{q}^{(k)} \leftarrow$ Equation (6.22) with $\mathbf{S}$ and $\mathbf{S}^{(k)}$
- 26: $b_v^{(k)} \leftarrow$ Equation (6.29) with $\mathbf{S}^{(k)}$ and b_{nom}
- 27: **end for**
- 28: $PL_v \leftarrow$ Equation (6.41) with $b_v^{(0)}, b_v^{(k)}, \mathbf{q}^{(k)}, T_k, \sigma_v^{(k)}, D_{k,v}, S_{v,k}, \forall k = 1, 2, \dots, N_{\text{fault modes}}$, and $f_{\varepsilon_j}, \forall j = 1, 2, \dots, n$
- 29: $\text{VPL} \leftarrow PL_3$
- 30: $\text{HPL} \leftarrow \sqrt{PL_1^2 + PL_2^2}$
- 31: No faults are detected
- 32: Return

and the non-Gaussian overbound $f_{\varrho,j,IF,acc}^i(x)$ are employed. For notations, the jack-knife ARAIM algorithm using the Gaussian overbound is named the JK-Gaussian ARAIM, while the one using the non-Gaussian overbound is named the JK-non-Gaussian ARAIM. Table 6.1 lists the usage of overbounds in different ARAIM algorithms in the experiment.

Table 6.1: Overbounds used in different ARAIM algorithms

Method	Baseline ARAIM	JK-Gaussian ARAIM	JK-non-Gaussian ARAIM
Overbounds	$f_{\varrho,j,IF,Gaussian}^i(x)$	$f_{\varrho,j,IF,Gaussian}^i(x)$	$f_{\varrho,j,IF,acc}^i(x)$

The integrity and continuity budget, the prior probability of satellite and constellation faults, and the threshold for the integrity risk coming from unmonitored faults are listed in Table 6.2. These values are aligned with the recommendation in the ARAIM algorithm description issued by WGC [100]. The maximum number of simultaneous faults ($k_{\max}$) that need to be monitored is determined by the method in [20], which is also given in Appendix A.2. For the single constellation case, $k_{\max} = 1$. For the dual constellation case, $k_{\max} = 2$. An equal allocation strategy is adopted in allocating the integrity and continuity budgets to each fault mode.

6.6.1 Single-constellation experiments

In this section, the performance of the proposed JK-Gaussian ARAIM and JK-non-Gaussian ARAIM algorithms is evaluated in the single GPS constellation setting, where the baseline ARAIM algorithm is taken as the benchmark. The first analysis involves the comparison between the baseline ARAIM and the proposed JK-Gaussian ARAIM algorithms, both of which use the Gaussian overbound for code IF combination nominal errors. Figures 6.1a and 6.1b show the map of 99.5 percentile of the VPL over the course of a day of the baseline ARAIM and the proposed JK-Gaussian

Table 6.2: Parameters used for evaluating integrity monitoring algorithms in the simulation

Parameter	Description	Value
I_{REQ}^3	Vertical integrity risk budget	9.8×10^{-8}
$I_{\text{REQ}}^1 + I_{\text{REQ}}^2$	Horizontal integrity risk budget	2×10^{-9}
$C_{\text{REQ,FA}}^3$	Vertical continuity risk budget allocated to false alarms	3.9×10^{-6}
$I_{\text{REQ}}^1 + I_{\text{REQ}}^2$	Horizontal continuity risk budget allocated to false alarms	9×10^{-8}
P_{sat}	Prior probability of satellite fault per approach	10^{-5}
P_{const}	Prior probability of constellation fault per approach	10^{-4}
P_{THRES}	Threshold for the integrity risk coming from unmonitored faults	9×10^{-8}

ARAIM algorithms, respectively. As can be seen, the two methods yield the same results, where the 99.5 percentile VPL is larger than 50 m in most user locations.

To gain a comprehensive understanding of the performance of the two methods, the triangular charts of the baseline ARAIM and the JK-Gaussian ARAIM regarding the vertical performance are plotted in Figure 6.1c and Figure 6.1d, respectively, which again demonstrates the equivalence of the two methods. Specifically, each bin in the triangular chart represents the number of occurrences of a specific pair of absolute vertical positioning error (VPE) and VPL among all 288×144 location-time events. The percentage of the normal operation (the VPL is larger than the VPE but less than the vertical alert limit (VAL), i.e., 35 m here) is around 86 %. The percentage of misleading information (the VPE is larger than the VPL but less than the VAL) and hazardously misleading information (the VPE is larger than the VAL without alerts) events are all zero for both methods.

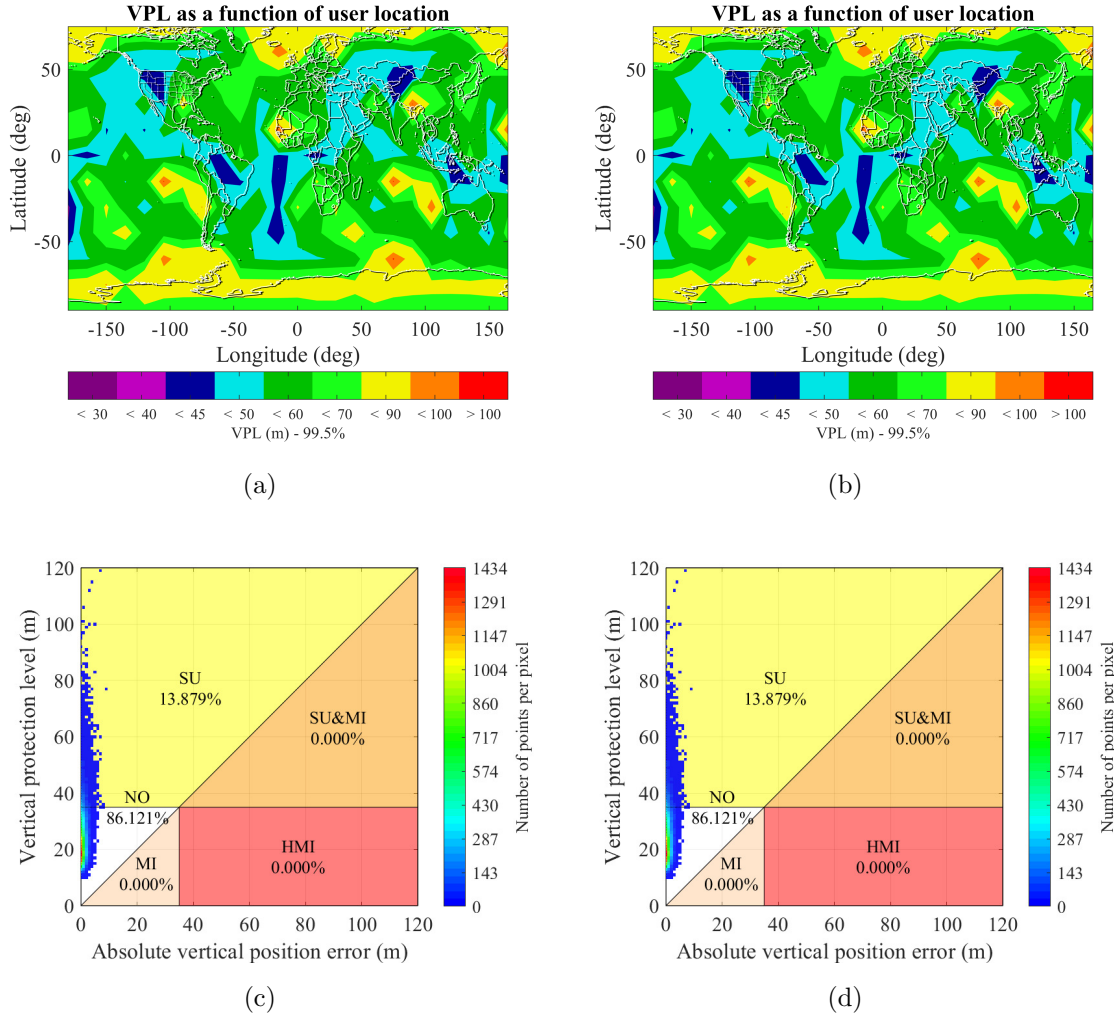


Figure 6.1: 99.5 percentile of the VPL over the course of the day yielded by (a) the baseline ARAIM and (b) the proposed JK-Gaussian ARAIM for the single constellation; and the triangular chart of (c) the baseline ARAIM and (d) the proposed JK-Gaussian ARAIM regarding the vertical performance for the single constellation. “NO” represents normal operation, “MI” represents misleading information, “SU” represents system unavailable, “SU&MI” represents system unavailable and misleading information, and “HMI” represents hazardously misleading information.

The second analysis focuses on the benefits brought by introducing non-Gaussian overbound into the jackknife ARAIM algorithm. Figure 6.2a shows the map of 99.5

percentile of the VPL over the course of a day of the proposed JK-non-Gaussian ARAIM algorithm. As can be seen, the 99.5 percentile VPL is less than 45 m in most user locations. By comparing to the results in Figure 6.1b, one can conclude that introducing non-Gaussian overbound into the jackknife ARAIM algorithm can reduce the VPL. The triangular chart of the JK-non-Gaussian ARAIM in Figure 6.2b further confirms this conclusion, where the distribution of the VPE-VPL pairs shows a higher concentration level than that of the jackknife ARAIM algorithm and the baseline ARAIM algorithm. More importantly, the percentage of the normal operation of the JK-non-Gaussian ARAIM method increases to 94.799 %, indicating that the JK-non-Gaussian ARAIM seldom comprises integrity.

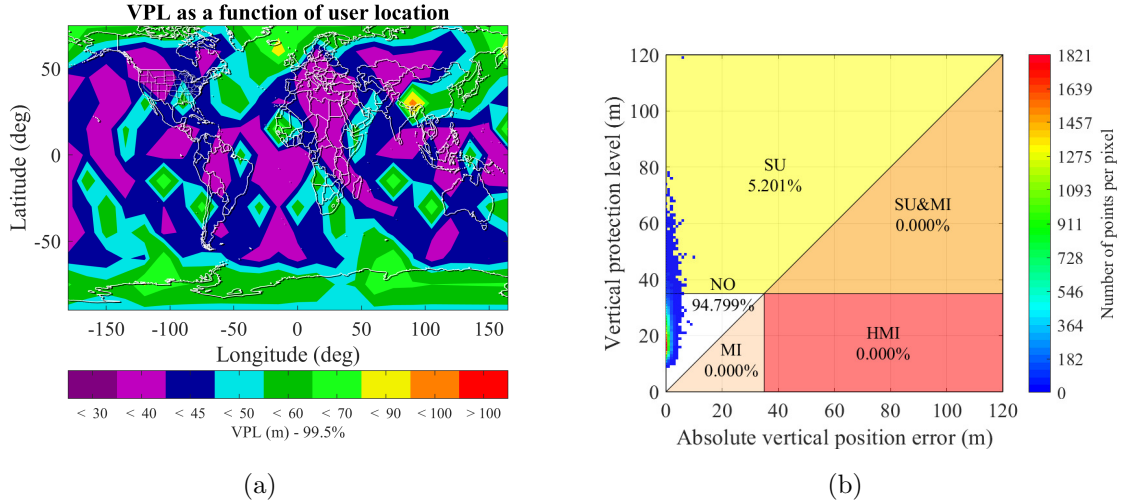


Figure 6.2: (a) 99.5 percentile of the VPL over the course of the day yielded by the proposed JK-non-Gaussian ARAIM for the single constellation; (b) The triangular chart of the proposed JK-non-Gaussian ARAIM regarding the vertical performance for the single constellation.

For a better understanding of the possibility of using the JK-non-Gaussian ARAIM to support LPV-200 precision approach operations, Table 6.3 summarizes the coverage of the three methods with $VAL = 35$ m at different levels of system availability. The system availability is the fraction of time that VPL is less than a given VAL at a given location, while the coverage is the fraction of the earth that satisfies a

given system availability. All the three methods show satisfactory performance in coverage under 75 % system availability. However, when the availability requirements increases to 95 %, the baseline ARAIM and the JK-Gaussian ARAIM algorithms only has a coverage of 15.16 %. In contrast, the coverage of the JK-non-Gaussian ARAIM still keeps above 88 % in this condition. Nevertheless, the coverage of the JK-non-Gaussian ARAIM decreases to 7.84 % under 99.5 % system availability. The above results reveal that the proposed JK-non-Gaussian ARAIM method has huge potential to support integrity applications with harsh navigation requirements.

Table 6.3: Coverage for the single constellation at different levels of system availability

VAL	Availability	Baseline ARAIM	JK-Gaussian ARAIM	JK-non-Gaussian ARAIM
35m	75 %	96.3 %	96.3 %	100 %
	95 %	15.16 %	15.16 %	88.64 %
	99.5 %	0 %	0 %	7.84 %

6.6.2 Dual-constellation experiments

This section evaluates the performance of the proposed JK-Gaussian ARAIM and JK-non-Gaussian ARAIM algorithms in the dual constellation setting. Similar to the single constellation setting in Section 5.3.2, the JK-Gaussian ARAIM exhibits the equivalent performance to the baseline ARAIM, as shown in the 99.5 percentile VPL map in Figures 6.3a and 6.3b. However, the magnitude of the 99.5 percentile VPL of these two methods exceeds 60 m at most user locations, which is significantly larger than that in the single constellation setting (see Figures 6.2a and 6.2b). This phenomenon shares the same reason for the observations in the multiply-fault detection experiment in Section 5.3.3: the SISRE of Galileo satellites in the dual constellation setting has significant heavy-tailed properties, which results in the over-conservatism

in the finalized Gaussian overbounds of code IF combination errors. Such conservatism is passed to the position domain bounding, eventually enlarging the VPLs in the dual-constellation setting. As a consequence, the system unavailability events of both methods experience a surge in the dual-constellation setting, which can be observed in the triangular chart in Figures 6.3c and 6.3d, where the system unavailability events with $VAL = 35$ m account for 45.674 %.

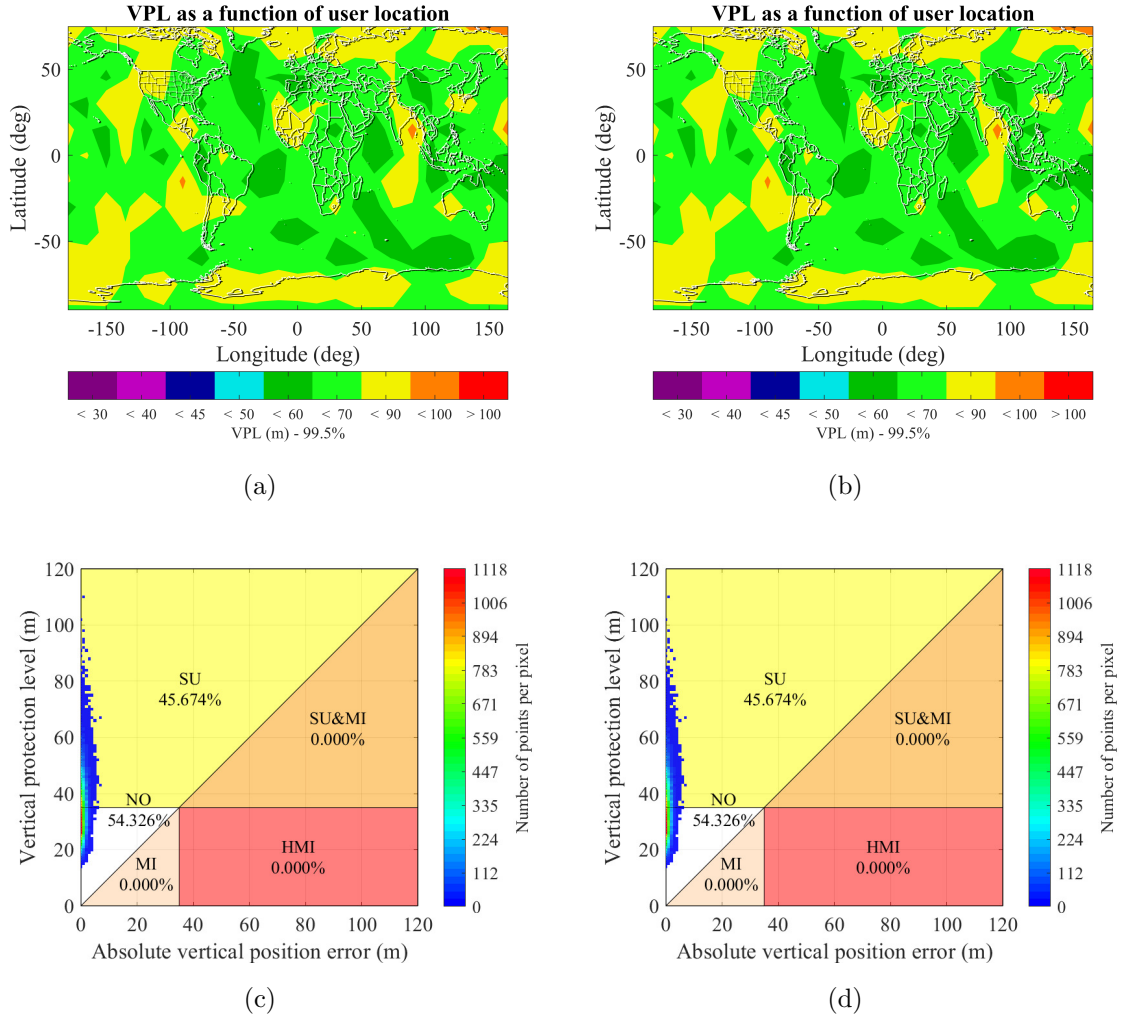


Figure 6.3: 99.5 percentile of the VPL over the course of the day yielded by (a) the baseline ARAIM and (b) the proposed JK-Gaussian ARAIM for the dual constellation; and the triangular chart of (c) the baseline ARAIM and (d) the proposed JK-Gaussian ARAIM regarding the vertical performance for the dual constellation.

Nevertheless, the JK-non-Gaussian ARAIM still shows satisfactory performance in the dual-constellation setting, where the 99.5 percentile VPL is smaller than 40 m in most user locations (Figure 6.4a) and the VPE-VPL pairs have extremely concentrated distribution (Figure 6.4b). Moreover, the percentage of the normal operation events with $VAL = 35$ m even exceeds 92 %, making it possible to support LPV-200 precision approach operations [27].

Table 6.4 summarizes the coverage of the three methods with $VAL = 35$ m at different levels of system availability. The baseline ARAIM and the JK-Gaussian ARAIM have a 54 % coverage even under 75 % system availability. This result is expected because both the baseline ARAIM and JK-Gaussian ARAIM use over-conservative Gaussian overbound. In contrast, the coverage of the JK-non-Gaussian ARAIM is nearly 100 % under 75 % system availability. Its coverage even exceed 62 % under 95 % system availability. These results reveal the huge potential of the JK-non-Gaussian ARAIM algorithm to support LPV-200 requirements using the GPS-Galileo dual constellation.

It is worth noting that the reporting result about the baseline ARAIM in this simulation study is quite different from the findings in [16, 84], from which the baseline ARAIM is examined to be able to provide global coverage for LPV-200 in GPS-Galileo dual constellation. The primary reason is that these studies use hypothetical models to simulate the range errors, which results in over-optimistic results. For example, the 1-sigma error bound of Galileo SISRE is set to be 0.96 m in [84], which is significantly smaller than the value determined by experimental data in Section 4.4. In such a condition, the system availability of baseline ARAIM is over-estimated.

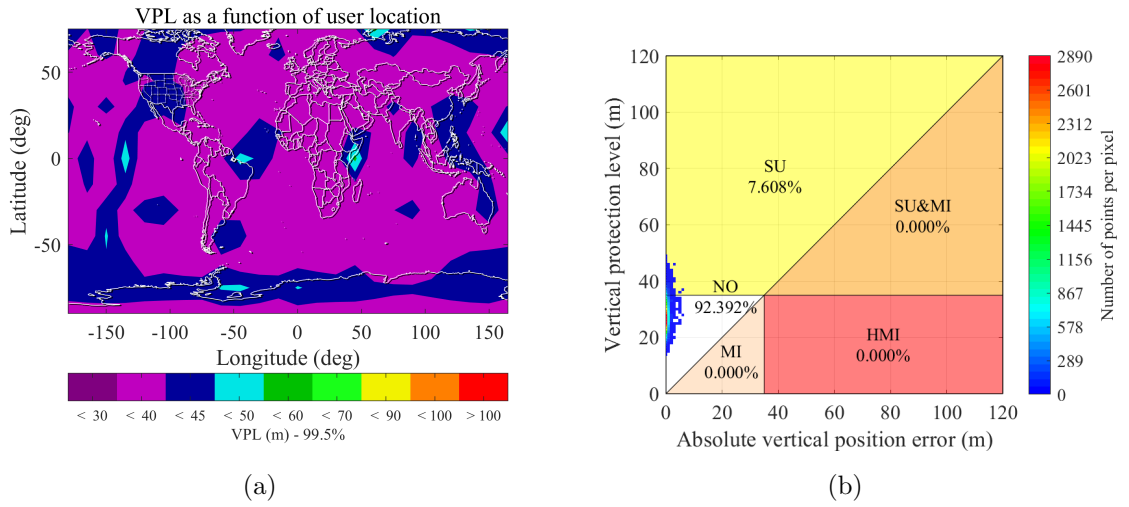


Figure 6.4: (a) 99.5 percentile of the VPL over the course of the day yielded by the proposed JK-non-Gaussian ARAIM for the dual constellation; (b) The triangular chart of the proposed JK-non-Gaussian ARAIM regarding the vertical performance for the dual constellation.

Table 6.4: Coverage for the dual constellation at different levels of system availability

VAL	Availability	Baseline ARAIM	JK-Gaussian ARAIM	JK-non-Gaussian ARAIM
35m	75 %	54 %	54 %	99.29 %
	95 %	0 %	0 %	62.55 %
	99.5 %	0 %	0 %	3.68 %

Chapter 7

Conclusion

This thesis develops the integrity monitoring algorithms with non-Gaussian nominal errors for safety-critical GNSS navigation in three aspects. These include the non-Gaussian overbounds for heavy-tailed SISRE, the fault detection method with non-Gaussian nominal errors, and the high-availability integrity monitoring algorithm. This thesis systematically introduces the non-Gaussian modeling to the integrity monitoring community and demonstrates its huge potential in improving system availability, shielding light on implementing integrity applications under stringent navigation requirements.

7.1 Summary of Accomplishments

(1) Nominal performance characterization of GPS/Galileo SISRE

This thesis details the methodology of determining satellite orbit and clock errors, including data source acquisition, antenna phase center offset correction for both broadcast and precise ephemerides, and computation of orbit and clock vectors in both ECEF and RAC frame, presenting a handy material for implementation. To support integrity applications, the UPE is calculated by projecting the orbit and

clock errors into earth grids within the satellite visibility cone and is regarded as the SISRE. The nominal performance of SISRE for each individual GPS and Galileo satellite during 2020–2022 has been analyzed, with emphasis put on the tail properties and the difference between GPS and Galileo satellites. The analyzed results provide the empirical basis for developing non-Gaussian overbounding methods and explaining the performance degradation of fault detection and integrity monitoring algorithms when incorporating Galileo constellations in the simulation studies.

(2) Non-Gaussian overbounding method for heavy-tailed errors

This thesis proposes the PGO for bounding heavy-tailed error distribution. Specifically, the BGMM is employed to fit the error distribution based on the EM algorithm. The CDF overbound of the BGMM is constructed based on the dominant relationship of each Gaussian component at the core and the tail region of the BGMM, respectively. In addition, a sigma inflation strategy is proposed to compensate the PGO, which allows it to bound the sample distribution as well as the fitted BGMM. The overbounding property is proven to be preserved through convolution, which makes it possible to derive pseudorange-level requirements from the position domain integrity requirements. To facilitate the range-to-position projection, a discretization technique that satisfies overbounding is further proposed and is used in the fast Fourier transform to reduce the computation time of convolution. The proposed PGO is applied to bound the GPS/Galileo SISRE, which provides the most competitive bounding performance when compared to the Gaussian overbound and Gaussian-Pareto overbound, yielding a sharp bound in both the core and tail parts of the error distribution. Based on the PGO of SISRE, the non-Gaussian nominal error bounds of code IF combination are further constructed, providing nominal error models for constructing non-Gaussian fault detection and integrity monitoring algorithms.

(3) Theoretical foundation for detecting faulty measurements under non-Gaussian nominal errors

This thesis develops a fault detection method with rigorous derivations for linearized pseudorange-based positioning systems under non-Gaussian nominal errors. Specifically, this thesis first constructs the jackknife test statistic by computing the inconsistency between the observed measurement and the predicted measurement based on subset solutions under the single-fault condition. Using the constructed jackknife test statistic, a jackknife detector is developed by formalizing a multiple-testing problem with the Bonferroni correction to detect faults. Then, the jackknife detector is extended to simultaneous faults by combining multiple test statistics. It is proven that the constructed test statistic is the linear combination of measurement errors without making assumptions about the distribution of errors, which provides an accurate probabilistic model for hypothesis testing and establishes theoretical foundations for fault detection.

(4) High-availability integrity monitoring algorithm under stringent navigation requirements

This thesis develops an integrity monitoring algorithm by systematically exploiting the properties of the jackknife detector in the range domain. Specifically, the proposed method constructs a set of hypotheses, including a fault-free hypothesis and alternative hypotheses, based on threat models. The jackknife detector is employed to construct the monitors, with its threshold determined by the allocated continuity budget. Then, the integrity risk under the fault-free hypothesis is bounded by ignoring knowledge of no detection. For alternative hypotheses, a tight bound of the integrity risk is derived by quantifying the impacts of hypothetical fault vectors on the position solution.

(5) Performance evaluation using real data in simulation

This thesis evaluates the performance of the proposed fault detection and integrity monitoring algorithm in a worldwide simulation. The range projection of clock and orbit error of the code IF combination is simulated based on authentic experimen-

tal data instead of relying on empirical models, which enhances the reliability of the experimental results. Results show that the proposed fault detection method demonstrates superior performance than the SS fault detection method under non-Gaussian nominal errors. Moreover, results reveal that the proposed non-Gaussian jackknife ARAIM algorithm has higher system availability than the baseline ARAIM method, making it possible to support LPV-200 requirements using the GPS-Galileo dual constellation.

7.2 Recommended Topics for Future Research

(1) Reduce conservatism of Galileo broadcast URA parameters

The analysis of SISRE for Galileo satellites reveals that the Galileo broadcast URA parameter is extremely conservative, which cannot fairly describe the nominal performance of Galileo SISRE at most times. As suggested by experiment results, such conservatism not only inflates the miss-detection rate of the fault detection algorithm but also enlarges the protection levels of integrity monitoring algorithms. It is recommended to use a sharper overbound than the Gaussian overbound in the Galileo broadcast information. The proposed PGO could be a good choice, but it is essential to balance the bounding performance and the communication cost introduced by additional bounding parameters. Using these non-Gaussian overbounds also poses challenges and brings opportunities for the design of the satellite broadcast message.

(2) Reduce computation cost by position-domain bounding

The proposed PGO provides a sharp yet conservative overbound for heavy-tailed errors in the range domain. However, the range-to-position projection of PGO does not have an analytical expression as that of the Gaussian overbound, thereby requiring convolution operations. This causes a computation burden in the proposed jackknife ARAIM algorithm, where the range-to-position projection needs to be executed in

each subset. Although a specialized discretization technique is proposed in this thesis and is used along with the fast Fourier transform to accelerate convolutions in the range-to-position projection, it still poses challenges for real-time integrity applications. Since the PGO has an analytical expression in PDF and CDF, it is possible to develop a parameterized overbound for the range-to-position projection of PGO. Such a position-domain overbound avoids the convolution operations, which enables the jackknife ARAIM algorithm to achieve the same computation efficiency as the baseline ARAIM algorithm, making it possible to realize real-time integrity monitoring with non-Gaussian nominal errors.

(3) Reduce conservatism of multi-testing in constructing monitors

Both the baseline ARAIM and the proposed jackknife ARAIM construct monitor and conduct hypothesis testing on each subset, which is a multi-testing problem. The Bonferroni correction is applied in both methods to handle this multi-testing problem. However, the Bonferroni procedure is known to be conservative, which can lead to a high miss-detection rate in the monitors. A possible remedy is to apply the Holm–Bonferroni correction [101], which keeps the family-wise error rate no higher than a pre-specified significance level. The Holm–Bonferroni correction shows less conservative in multi-testing problems than the Bonferroni correction. However, Holm–Bonferroni correction involves the systematic adjustment of the significance level for each individual test. It is essential to investigate and remove the impacts of such adjustments on system integrity.

(4) Integrity risk evaluation with fault exclusion

The proposed integrity monitoring algorithm warns the users once a fault is detected. Future research can develop fault exclusion algorithms to further increase the system availability. A heuristic method is to rank the p-value of the test statistics on each subset and exclude the one with the smallest p-value. However, due to the masking effects of simultaneous faults, it is possible to exclude nominal measurements. There-

fore, special caution should be given in evaluating the impacts of wrong exclusions on the integrity risk. Additional tests should be constructed to justify the wrong exclusion probability, which should be further considered in the PL calculation.

(5) Incorporate additional constellations to improve availability

In the GPS-Galileo dual constellation experiment, the proposed JK-Gaussian ARAIM and the baseline ARAIM algorithms exhibit significant degradation in system availability due to the heavy-tailed properties of Galileo SISRE. To further improve the system availability, one can incorporate additional constellations, such as BeiDou, GLONASS, and QZSS constellations. The clock and orbit errors of these constellations should be carefully computed by employing suitable corrections, such as the antenna phase center offset correction mentioned in Chapter 3, the time group delay correction, and the clock corrections.

7.3 Closing

The non-Gaussian integrity monitoring algorithm presented in this thesis demonstrates high availability under stringent navigation requirements, making it possible to achieve the LPV-200 requirements using the recent GPS-Galileo dual constellation.

Appendix A

Appendices

A.1 Gaussian Overbound of Multipath and Code Noise

The Gaussian overbound for multipath and code noise error for code IF combination is given by

$$\sigma_{\varrho, user, j, AB}^i = \sigma_{\varrho, user, j}^i \sqrt{\frac{\gamma^2 + 1}{(\gamma - 1)^2}}, \quad (\text{A.1})$$

where γ is the ratio of squares of two frequencies, as defined in Equation (2.9), and

$$\sigma_{\varrho, user, j}^i = \sqrt{(\sigma_{\varrho, noise, j}^i)^2 + (\sigma_{\varrho, multipath, j}^i)^2}. \quad (\text{A.2})$$

The code noise bound $\sigma_{\varrho, noise, j}^i$ and multipath bound $\sigma_{\varrho, multipath, j}^i$ after carrier smoothing suggested by WGC are provided by

(1) GPS Airborne Receiver [102]

$$\sigma_{\varrho, noise, j}^i = 0.15[m] + 0.43[m] \exp\left(-\frac{\theta_j^i[deg]}{6.9}\right) \quad (\text{A.3a})$$

$$\sigma_{\varrho, multipath, j}^i = 0.13[m] + 0.53[m] \exp\left(-\frac{\theta_j^i[deg]}{10}\right). \quad (\text{A.3b})$$

(2) Galileo Airborne Receiver [41]

Table A.1: The code noise and multipath error bound for Galileo airborne receiver against the elevation angle

$\theta_j^i[deg]$	$\sigma_{\varrho,user,j}^i$	$\theta_j^i[deg]$	$\sigma_{\varrho,user,j}^i$	$\theta_j^i[deg]$	$\sigma_{\varrho,user,j}^i$
5	0.4529	35	0.2504	65	0.2295
10	0.3553	40	0.2438	70	0.2278
15	0.3063	45	0.2396	75	0.2297
20	0.2638	50	0.2359	80	0.2310
25	0.2593	55	0.2339	85	0.2274
30	0.2555	60	0.2302	90	0.2277

A.2 Maximum Number of Simultaneous Faults That Need to be Monitored

This section describes the procedure of determining the maximum number of simultaneous faults that need to be monitored $k_{\max}$ and the prior probabilities of fault mode P_{H_k} in the baseline ARAIM algorithm [20].

Define

$$P_{event,i} = P_{sat,i} \tag{A.4a}$$

$$P_{event,N_{sat}+j} = P_{const,j} . \tag{A.4b}$$

Assume fault mode k includes the events in set Ω_{H_i} . Then the prior probabilities of fault mode P_{H_k} is defined as

$$P_{H_k} = \prod_{j \in \Omega_{H_i}} P_{event,j} . \tag{A.5}$$

The probability of all fault modes composed of r or more independent events is given by

$$P_{>}(r) = \sum_{k=r}^{N_{sat}+N_{const}} \sum_{i_1 < i_2 < \dots < i_k} \prod_{j \in i_1, i_2, \dots, i_k} P_{event,j} \prod_{j \notin i_1, i_2, \dots, i_k} (1 - P_{event,j}). \quad (\text{A.6})$$

It is proven in [20] that

$$P_{>}(r) \leq \sum_{i_1 < i_2 < \dots < i_r} P_{event,i_1} P_{event,i_2} \dots P_{event,i_r}. \quad (\text{A.7})$$

Since

$$\left(\sum_{k=1}^{N_{sat}+N_{const}} P_{event,k} \right)^r \geq \sum_{i_1 \neq i_2 \neq \dots \neq i_r} P_{event,i_1} P_{event,i_2} \dots P_{event,i_r} \quad (\text{A.8a})$$

$$= r! \sum_{i_1 < i_2 < \dots < i_r} P_{event,i_1} P_{event,i_2} \dots P_{event,i_r}, \quad (\text{A.8b})$$

where Equation (A.8b) indicates that there are $r!$ ways to arrange a sequence of length r , $P(r)$ in Equation (A.7) can be bounded by

$$P_{>}(r) \leq \frac{\left(\sum_{k=1}^{N_{sat}+N_{const}} P_{event,k} \right)^r}{r!}. \quad (\text{A.9})$$

Therefore, $k_{\max}$ is determined by

$$k_{\max} = \varphi_{P_{THRES}} \left(\sum \sum_{k=1}^{N_{sat}+N_{const}} P_{event,k} \right), \quad (\text{A.10})$$

where

$$\varphi_{P_{THRES}}(u) = \min \left\{ r \mid \frac{u^{r+1}}{(r+1)!} \leq P_{THRES} \right\}, \quad (\text{A.11})$$

and P_{THRES} is the threshold for the integrity risk coming from unmonitored faults. Then, the probability of the unmonitored events is given by

$$P_{\text{not monitored}} = P_{>}(k_{\max} + 1). \quad (\text{A.12})$$

A.3 The EM Algorithm

The PDF of a K -component GMM model with mixture weights of $p_1, p_2, \dots, p_K$ can be written as

$$f(x|\Theta) = \sum_{k=1}^K p_k f_N(x; b_k, \sigma_k), \quad (\text{A.13})$$

where Θ represents all parameters, including each Gaussian component's mixture weight, mean, and variance. For a given observation x_i , we define the allocation (latent) variable $c = \{1, 2, \dots, K\}$ that marks the Gaussian component from which x_i is generated. Then, the mixture weight of Gaussian components can be interpreted as the prior probability of the allocation variable as follows,

$$p_k = P(c = k). \quad (\text{A.14})$$

The membership weight of the GMM can be defined based on Bayes' Theorem as follows,

$$w_{i,k} = P(c = k|x_i) = \frac{p_k f_N(x_i; b_k, \sigma_k)}{\sum_{k=1}^K p_k f_N(x_i; b_k, \sigma_k)}, \quad (\text{A.15})$$

which indicates the posterior probability of a data point x_i being generated from the k th Gaussian component. The log-likelihood function can be written as

$$L(x|\Theta) = \sum_{i=1}^N \ln \sum_{k=1}^K P(x_i|c = k, b_k, \sigma_k) P(c = k). \quad (\text{A.16})$$

By utilizing Jensen's inequality [103], the above equation can be simplified as

$$L(x|\Theta) \geq Q(x|\Theta), \quad (\text{A.17a})$$

$$\begin{aligned} Q(x|\Theta) &= \sum_{i=1}^N \sum_{k=1}^K w_{i,k} \ln \frac{p_k f_N(x_i; b_k, \sigma_k)}{w_{i,k}} \\ &= \sum_{i=1}^N \sum_{k=1}^K w_{i,k} \left(\ln p_k - \frac{(x_i - b_k)^2}{2\sigma_k^2} - \ln w_{i,k} - \ln \sqrt{2\pi\sigma_k^2} \right), \end{aligned} \quad (\text{A.17b})$$

where $Q(x|\Theta)$ is the lower bound of $L(x|\Theta)$. Therefore, the model parameters of GMM can be estimated by maximizing $Q(x|\Theta)$, which can be effectively achieved

by the EM algorithm [74]. The EM algorithm is an iterative supervised training algorithm consisting of two steps: the E-step and the M-step. Firstly, the model parameters are randomly initialized. In the E-step, the algorithm calculates the value of membership weight $w_{i,k}$ based on the latest model parameters, as shown in (A.15). In the M-step, the algorithm updates the model parameters based on the value of $w_{i,k}$. In this step, the log-likelihood function is maximized by taking the partial derivative over the model parameters Θ and letting the partial derivative equal to zero. The two steps are repeated until convergence is reached.

A.4 Proof of Tail Region Overbound

Recall that Section 4.3.4.3.3 constructs the overbound distribution in the left-tail region as follows:

$$G_o^L(x) = (1 + k) (1 - p_1) G(x; 0, \sigma_2) \quad \forall x < x_{lp}. \quad (\text{A.18})$$

The difference between $G_o^L(x)$ and $G(x)$ is given by

$$G_o^L(x) - G(x) = k (1 - p_1) G(x; 0, \sigma_2) - p_1 G(x; 0, \sigma_1). \quad (\text{A.19})$$

By substituting (4.28) into (A.19),

$$\begin{aligned} G_o^L(x) - G(x) = \frac{p_1}{G(x_{lp}; 0, \sigma_2)} & \left(G(x_{lp}; 0, \sigma_1) G(x; 0, \sigma_2) \right. \\ & \left. - G(x; 0, \sigma_1) G(x_{lp}; 0, \sigma_2) \right). \end{aligned} \quad (\text{A.20})$$

If $G_o^L(x)$ is the CDF overbound of $G(x)$, then the right-hand side of (A.20) should be non-negative, which is equivalent to prove

$$\frac{G(x_{lp}; 0, \sigma_1)}{G(x_{lp}; 0, \sigma_2)} > \frac{G(x; 0, \sigma_1)}{G(x; 0, \sigma_2)} \quad \forall x < x_{lp}. \quad (\text{A.21})$$

Let $f(x) = \frac{G(x;0,\sigma_1)}{G(x;0,\sigma_2)}$, then the derivative of $f(x)$ is given by

$$f'(x) = \frac{f_N(x;0,\sigma_1)G(x;0,\sigma_2) - f_N(x;0,\sigma_2)G(x;0,\sigma_1)}{G^2(x;0,\sigma_2)}. \quad (\text{A.22})$$

Define

$$m(\sigma; x) = \frac{f_N(x;0,\sigma)}{G(x;0,\sigma)} = \frac{\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{x^2}{2\sigma^2}\right)}{\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x}{\sigma\sqrt{2}}\right)\right]} \quad (\text{A.23a})$$

$$\forall \sigma > 0, x < x_{lp},$$

$$h(\gamma; x) = m\left(-\frac{1}{\gamma}; x\right) \quad \forall \gamma < 0, x < x_{lp}. \quad (\text{A.23b})$$

Then $h(\gamma; x)$ can be written as

$$h(\gamma; x) = \frac{\frac{\gamma}{-\sqrt{2\pi}}\exp\left(-\frac{\gamma^2}{2\left(-\frac{1}{x}\right)^2}\right)}{\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\gamma}{\left(-\frac{1}{x}\right)\sqrt{2}}\right)\right]} = \frac{\frac{\gamma}{x} \cdot \frac{1}{\left(-\frac{1}{x}\right)\sqrt{2\pi}}\exp\left(-\frac{\gamma^2}{2\left(-\frac{1}{x}\right)^2}\right)}{\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\gamma}{\left(-\frac{1}{x}\right)\sqrt{2}}\right)\right]} = \frac{1}{x} \cdot \gamma \frac{f_N\left(\gamma; 0, \left(-\frac{1}{x}\right)^2\right)}{G\left(\gamma; 0, -\frac{1}{x}\right)}. \quad (\text{A.24})$$

Define $\lambda(\gamma) = \frac{f_N(\gamma; 0, (-\frac{1}{x}))}{G(\gamma; 0, -\frac{1}{x})}$, which is known as the inverse Mills ratio [104] and has the following properties [105]:

$$\lambda(\gamma) \geq 0 \quad \forall \gamma \quad (\text{A.25a})$$

$$\lambda'(\gamma) < 0 \quad \forall \gamma. \quad (\text{A.25b})$$

Then the partial derivative of $h(\gamma; x)$ with respect to γ can be calculated as

$$\frac{\partial h(\gamma; x)}{\partial \gamma} = \frac{1}{x} \left[\lambda(\gamma) + \gamma \lambda'(\gamma) \right]. \quad (\text{A.26})$$

According to the chains rule,

$$\begin{aligned} \frac{\partial m(\sigma; x)}{\partial \sigma} &= \frac{\partial h\left(-\frac{1}{\sigma}; x\right)}{\partial \sigma} = \frac{\partial h\left(-\frac{1}{\sigma}; x\right)}{\partial \gamma} \cdot \frac{\partial\left(-\frac{1}{\sigma}\right)}{\partial \sigma} \\ &= \frac{1}{x} \left[\lambda\left(-\frac{1}{\sigma}\right) - \frac{1}{\sigma} \lambda'\left(-\frac{1}{\sigma}\right) \right] \frac{1}{\sigma^2}. \end{aligned} \quad (\text{A.27})$$

When $x < 0$, we can easily prove that $\frac{\partial m(\sigma; x)}{\partial \sigma} < 0$. This means that $m(\sigma; x)$ is a monotonically decreasing function when $x < 0$. Therefore,

$$\begin{aligned} m(\sigma_1; x) &= \frac{f_N(x; 0, \sigma_1)}{G(x; 0, \sigma_1)} > \frac{f_N(x; 0, \sigma_2)}{G(x; 0, \sigma_2)} \\ &= m(\sigma_2; x) \quad \forall \sigma_2 > \sigma_1 > 0, x < 0. \end{aligned} \quad (\text{A.28})$$

This is equivalent to saying that (A.22) has the property as follows:

$$f'(x) > 0 \quad \forall \sigma_2 > \sigma_1 > 0, x < 0. \quad (\text{A.29})$$

Equation (A.29) indicates that $f(x)$ is a monotonically increasing function when $x < 0$, and therefore (A.21) is proved. Then, we can conclude that $G_o^L(x)$ is the CDF overbound of $G(x)$ when $x < x_{lp}$. According to the symmetric property of $G_o^L(x)$ and $G_o^R(x)$, $G_o^R(x)$ is easy to be proved as the CDF overbound of $G(x)$ when $x > x_{rp}$. This ends the proof.

A.5 Proof of Monotonicity

In the PGO, its PDF discontinues at the core-tail transition points x_{lp} and x_{rp} . The value of the leap at x_{lp} can be written as

$$f_o^{\text{core}}(x_{lp}) - f_o^L(x_{lp}) = -(1+k)(1-p_1)f_N(x_{lp}; 0, \sigma_2) + p_1f_N(x_{lp}; 0, \sigma_1) + c. \quad (\text{A.30})$$

Define

$$k^* = \frac{p_1f_N(x_{lp}; 0, \sigma_1)}{(1-p_1)f_N(x_{lp}; 0, \sigma_2)}. \quad (\text{A.31})$$

As proved in (A.28) in Appendix A.4,

$$\frac{f_N(x; 0, \sigma_1)}{G(x; 0, \sigma_1)} > \frac{f_N(x; 0, \sigma_2)}{G(x; 0, \sigma_2)} \quad \forall \sigma_2 > \sigma_1 > 0, x < 0. \quad (\text{A.32})$$

We have

$$k^* = \frac{p_1f_N(x_{lp}; 0, \sigma_1)}{(1-p_1)f_N(x_{lp}; 0, \sigma_2)} > \frac{p_1G(x_{lp}; 0, \sigma_1)}{(1-p_1)G(x_{lp}; 0, \sigma_2)} = k > 0 \quad \forall \sigma_2 > \sigma_1 > 0, x < 0. \quad (\text{A.33})$$

Therefore,

$$f_o^{\text{core}}(x_{lp}) - f_o^L(x_{lp}) > -(1 + k^*)(1 - p_1)f_N(x_{lp}; 0, \sigma_2) + p_1f_N(x_{lp}; 0, \sigma_1) + c. \quad (\text{A.34})$$

Substituting (4.33) and (A.31) into (A.34), we have

$$f_o^{\text{core}}(x_{lp}) - f_o^L(x_{lp}) > \frac{1 - p_1}{x_{lp}} (G(x_{lp}; 0, \sigma_2) - x_{lp}f_N(x_{lp}; 0, \sigma_2) - 0.5). \quad (\text{A.35})$$

Define

$$g(x_{lp}) = G(x_{lp}; 0, \sigma_2) - x_{lp}f_N(x_{lp}; 0, \sigma_2) - 0.5, \quad (\text{A.36})$$

then its first derivative can be calculated as

$$g'(x_{lp}) = \left(1 + \frac{x_{lp}^2}{\sigma_2^2}\right) f_N(x_{lp}; 0, \sigma_2) > 0. \quad (\text{A.37})$$

Since $g(0) = 0$, we have

$$g(x_{lp}) < 0 \quad \forall x_{lp} < 0. \quad (\text{A.38})$$

Therefore, (A.35) can be written as

$$f_o^{\text{core}}(x_{lp}) - f_o^L(x_{lp}) > \frac{1 - p_1}{x_{lp}} g(x_{lp}) > 0 \quad \forall x_{lp} < 0. \quad (\text{A.39})$$

Since

$$f_o^{L'}(x) = -\frac{x}{\sigma_2^2} (1 + k) p_2 f_N(x; 0, \sigma_2) > 0 \quad \forall x < x_{lp} \quad (\text{A.40a})$$

$$f_o^{\text{core}'}(x) = -\frac{x}{\sigma_1^2} p_1 f_N(x; 0, \sigma_1) > 0 \quad \forall x_{lp} < x < 0, \quad (\text{A.40b})$$

we can conclude that $f_o(x)$ is a monotonically increasing function when $x < 0$.

A.6 Distribution of Linear Combination of Random Variables

Consider a set of independent random variables $Y_1, Y_2, \dots, Y_N$ with zero means. The PDF of Y_i is given by $f_{Y_i}(x)$. Let Y be the linear combination of $Y_1, Y_2, \dots, Y_N$ as

follows:

$$Y = \sum_{i=1}^N a_i Y_i, \quad (\text{A.41})$$

where $a_i, i = 1, 2, \dots, N$ is the coefficient. Let

$$\tilde{Y}_i = a_i Y_i, \quad i = 1, 2, \dots, N \quad (\text{A.42})$$

be a new random variable, and then Y can be written by

$$Y = \sum_{i=1}^N \tilde{Y}_i. \quad (\text{A.43})$$

The CDF of $\tilde{Y}_i$ is given by

$$F_{\tilde{Y}_i}(y) = P(\tilde{Y}_i < y) = P\left(Y_i < \frac{y}{|a_i|}\right) = \int_{-\infty}^{y/|a_i|} f_{Y_i}(x) dx. \quad (\text{A.44})$$

Let $t = |a_i| x$, then we have

$$F_{\tilde{Y}_i}(y) = \frac{1}{|a_i|} \int_{-\infty}^y f_{Y_i}\left(\frac{t}{|a_i|}\right) dt. \quad (\text{A.45})$$

Therefore, the PDF of $\tilde{Y}_i$ is given by

$$f_{\tilde{Y}_i}(x) = \frac{1}{|a_i|} f_{Y_i}\left(\frac{x}{|a_i|}\right). \quad (\text{A.46})$$

Finally, the PDF of Y can be obtained through convolution as

$$\begin{aligned} f_Y(x) &= f_{\tilde{Y}_1}(x) * f_{\tilde{Y}_2}(x) * \dots * f_{\tilde{Y}_N}(x) \\ &= \prod_{i=1}^n |a_i|^{-1} f_{Y_1}\left(\frac{x}{|a_1|}\right) * f_{Y_2}\left(\frac{x}{|a_2|}\right) * \dots * f_{Y_n}\left(\frac{x}{|a_n|}\right). \end{aligned} \quad (\text{A.47})$$

A.7 Distribution of Jackknife Residual under Gaussian Noises

The Gauss-Markov conditions concern the set of noises in the linear system $y = \mathbf{G}x + \boldsymbol{\varepsilon}$ as follows:

1. Zero mean: $E[\varepsilon_i] = 0 \forall i$;
2. Homoscedastic: $\text{Var}[\varepsilon_i] = \sigma^2 < \infty \forall i$;
3. Uncorrelated: $\text{Cov}[\varepsilon_i, \varepsilon_j] = 0 \forall i \neq j$.

Under Gauss-Markov conditions, the ordinary least squares (OLS) estimator is the best linear unbiased estimator (BLUE).

A further generalization of the Gauss-Markov conditions to heteroscedastic and correlated errors has been developed [106], and its application to the weighted least squares (WLS) estimator can be stated as follows:

“WLS is the BLUE if the weight matrix is equal to the inverse of the variance-covariance matrix of the noises.”

Based on the generalized Gauss-Markov conditions, the subsolution in Equation (2.38a) has the following properties:

$$E[\hat{\mathbf{x}}^{(k)}] = 0 \quad (\text{A.48a})$$

$$\text{Var}[\hat{\mathbf{x}}^{(k)}] = \mathbf{S}^{(k)} \mathbf{W}^{-1} \mathbf{S}^{(k)T} . \quad (\text{A.48b})$$

By substituting Equation (5.9) into Equation (5.10), the jackknife residual can be written by

$$\begin{aligned} t_k &= y_k - \mathbf{g}_k \hat{\mathbf{x}}^{(k)} \\ &= \mathbf{g}_k \mathbf{x}^{(k)} + \varepsilon_k - \mathbf{g}_k \hat{\mathbf{x}}^{(k)} \\ &= \mathbf{g}_k (\mathbf{x}^{(k)} - \hat{\mathbf{x}}^{(k)}) + \varepsilon_k . \end{aligned} \quad (\text{A.49})$$

The expectation and variance of the jackknife residual in (A.49) are given by

$$E[t_k] = 0 \quad (\text{A.50a})$$

$$\begin{aligned} \text{Var}[t_k] &= \mathbf{g}_k \text{Var} [\mathbf{x}^{(k)} - \hat{\mathbf{x}}^{(k)}] \mathbf{g}_k^T + \sigma_k^2 \\ &= \mathbf{g}_k \mathbf{S}^{(k)} \mathbf{W}^{-1} \mathbf{S}^{(k)T} \mathbf{g}_k^T + \sigma_k^2 . \end{aligned} \quad (\text{A.50b})$$

As shown in Equation (5.19), the Jackknife residual can be rewritten as

$$t_k = \tilde{\mathbf{p}}_k \boldsymbol{\varepsilon}, \quad (\text{A.51})$$

which is a linear combination of measurement noises. If ε_i has a zero-mean Gaussian distribution defined in Equation (5.22), t_k will have a Gaussian distribution

$$t_k \sim \mathcal{N} \left(0, \tilde{\mathbf{p}}_k \mathbf{W}^{-1} \tilde{\mathbf{p}}_k^T \right). \quad (\text{A.52})$$

Since a Gaussian distribution is uniquely defined by its mean and variance, the following equation will hold:

$$\mathbf{g}_k \mathbf{S}^{(k)} \mathbf{W}^{-1} \mathbf{S}^{(k)T} \mathbf{g}_k^T + \sigma_k^2 = \tilde{\mathbf{p}}_k \mathbf{W}^{-1} \tilde{\mathbf{p}}_k^T. \quad (\text{A.53})$$

Therefore,

$$t_k \sim \mathcal{N} \left(0, \mathbf{g}_k \mathbf{S}^{(k)} \mathbf{W}^{-1} \mathbf{S}^{(k)T} \mathbf{g}_k^T + \sigma_k^2 \right). \quad (\text{A.54})$$

A.8 Bonferroni Correction

The hypotheses with Bonferroni correction [93] in Equation (5.25) have the following relationship with the original hypotheses in Equation (5.24):

$$\begin{aligned} H_0 &= \bigcap_{k=1}^n H_0^{(k)} \\ H_1 &= \bigcup_{k=1}^n H_1^{(k)}. \end{aligned} \quad (\text{A.55})$$

Assume that the probability of type I error of the corrected hypothesis test is α^* . Then,

$$\begin{aligned}
1 - \alpha^* &= P(\text{All tests accept} \mid H_0) \\
&= 1 - P(\text{At least one test is rejected} \mid H_0) \\
&\geq 1 - \sum_{k=1}^n P(\text{Origin test } i \text{ is rejected} \mid H_0) \\
&= 1 - \sum_{k=1}^n P(\text{Origin test } i \text{ is rejected} \mid H_0^{(k)}) \\
&= 1 - n\alpha .
\end{aligned} \tag{A.56}$$

In addition,

$$\begin{aligned}
\alpha^* &= P(\text{At least one test is rejected} \mid H_0) \\
&\geq P(\text{Origin test } i \text{ is rejected} \mid H_0^{(k)}) \\
&= \alpha .
\end{aligned} \tag{A.57}$$

Therefore,

$$\alpha \leq \alpha^* \leq n\alpha . \tag{A.58}$$

To keep the type I error α^* not exceeding τ (e.g., 0.05),

$$n\alpha = \tau . \tag{A.59}$$

Thus, the type I error of the individual test would be $\alpha = \frac{\tau}{n}$.

Bibliography

- [1] R. Wang and T. Walter, “Characterization and comparison of Galileo and GPS anomalies,” in *Proceedings of the 2023 International Technical Meeting of The Institute of Navigation*, 2023, pp. 597–610.
- [2] T. Walter and P. Enge, “Weighted RAIM for precision approach,” in *Proceedings of the 8th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1995)*, vol. 8, no. 1. Institute of Navigation, 1995, pp. 1995–2004.
- [3] R. G. Brown, “A baseline GPS RAIM scheme and a note on the equivalence of three RAIM methods,” *NAVIGATION*, vol. 39, no. 3, pp. 301–316, Sep. 1992. [Online]. Available: <https://onlinelibrary.wiley.com/doi/10.1002/j.2161-4296.1992.tb02278.x>
- [4] Y. C. Lee, “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 (1986)*, 1986, pp. 1–4.
- [5] A. K. Brown, “Receiver autonomous integrity monitoring using a 24-satellite GPS constellation,” in *Proceedings of the Satellite Division’s First Technical Meeting (ION GPS 1987)*, vol. 1, 1987, pp. 256–262.
- [6] R. G. Brown and P. W. McBurney, “Self-contained GPS integrity check using maximum solution separation,” *NAVIGATION*, vol. 35, no. 1, pp. 41–53, 1988.

- [7] A. K. Brown, “Civil aviation integrity requirements for the global positioning system,” *NAVIGATION*, vol. 35, no. 1, pp. 23–40, 1988.
- [8] B. W. Parkinson and P. Axelrad, “Autonomous GPS integrity monitoring using the pseudorange residual,” *NAVIGATION*, vol. 35, no. 2, pp. 255–274, Jun. 1988.
- [9] M. A. Sturza, “Navigation system integrity monitoring using redundant measurements,” *NAVIGATION*, vol. 35, no. 4, pp. 483–501, 1988.
- [10] R. G. Brown, G. Y. Chin, and J. H. Kraemer, “RAIM: Will it meet the RTCA GPS minimum operational performance standards?” in *Proceedings of the 1991 National Technical Meeting of The Institute of Navigation*, 1991, pp. 103–111.
- [11] R. Fontana and D. Latterman, “GPS modernization and the future,” in *Proceedings of the IAIN World Congress and the 56th Annual Meeting of The Institute of Navigation (2000)*, 2000, pp. 222–231.
- [12] P. Enge, “GPS modernization: capabilities of the new civil signals,” in *Australian International Aerospace Congress*, vol. 29, 2003.
- [13] R. Haddad, K. Kovach, R. Slattery, and J. Gillis, “GPS modernization and beyond,” in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. IEEE, 2020, pp. 399–406.
- [14] F. A. Administration, “Phase II of the GNSS Evolutionary Architecture Study,” 2010, accessed on April 2024. [Online]. Available: https://www.faa.gov/sites/faa.gov/files/about/office_org/headquarters_offices/ato/GEASPhaseII_Final.pdf
- [15] B. S. Pervan, S. P. Pullen, and J. R. Christie, “A multiple hypothesis approach to satellite navigation integrity,” *NAVIGATION*, vol. 45, no. 1, pp. 61–71, 1998.

- [16] J. Blanch, T. Walter, and P. Enge, “RAIM with optimal integrity and continuity allocations under multiple failures,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 3, pp. 1235–1247, 2010.
- [17] M. Joerger, F.-C. Chan, S. Langel, and B. Pervan, “RAIM detector and estimator design to minimize the integrity risk,” in *Proceedings of the 25th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2012)*, 2012, pp. 2785–2807.
- [18] M. Joerger, F.-C. Chan, and B. Pervan, “Solution separation versus residual-based RAIM,” *NAVIGATION*, vol. 61, no. 4, pp. 273–291, Dec. 2014.
- [19] J. Blanch, T. Walter, P. Enge, Y. Lee, B. Pervan, M. Rippl, and A. Spletter, “Advanced RAIM user algorithm description: Integrity support message processing, fault detection, exclusion, and protection level calculation,” in *Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012)*, Nashville, TN, Sep. 2012, pp. 2828–2849.
- [20] J. Blanch, T. Walker, P. Enge, Y. Lee, B. Pervan, M. Rippl, A. Spletter, and V. Kropp, “Baseline advanced RAIM user algorithm and possible improvements,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 713–732, 2015.
- [21] J. Blanch, T. Walter, C. Milner, M. Joerger, B. Pervan, and D. Bouvet, “Baseline advanced RAIM user algorithm: Proposed updates,” in *Proceedings of the 2022 International Technical Meeting of The Institute of Navigation*, 2022, pp. 229–251.
- [22] J. Rife, S. Pullen, and B. Pervan, “Core overbounding and its implications for LAAS integrity,” in *Proceedings of the 17th International Technical Meeting of*

- the Satellite Division of The Institute of Navigation (ION GNSS 2004)*, Long Beach, CA, Sep. 2004, pp. 2810–2821.
- [23] R. Braff and C. Shively, “A method of over bounding ground based augmentation system (GBAS) heavy tail error distributions,” *The Journal of Navigation*, vol. 58, no. 1, pp. 83–103, 2005.
- [24] B. Pervan, S. Pullen, and I. Sayim, “Sigma estimation, inflation, and monitoring in the LAAS ground system,” in *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, 2000, pp. 1234–1244.
- [25] S. Perea, M. Meurer, M. Rippl, B. Belabbas, and M. Joerger, “URA/SISA analysis for GPS and Galileo to support ARAIM,” *NAVIGATION*, vol. 64, no. 2, pp. 237–254, 2017.
- [26] S. Wang, Y. Zhai, and X. Zhan, “Characterizing bds signal-in-space performance from integrity perspective,” *NAVIGATION*, vol. 68, no. 1, pp. 157–183, 2021.
- [27] International Civil Aviation Organisation, “Annex 10, Aeronautical Telecommunications, Volume I (Radio Navigational Aids),” July 2006.
- [28] E. U. A. S. Agency, “NPA 2024-06: Introduction of a regulatory framework for the operation of drones — Enabling the initial airworthiness of UAS subject to certification, and the continuing airworthiness of those UAS operated in the ‘specific’ category,” 2024, accessed on April 2025. [Online]. Available: <https://www.easa.europa.eu/en/document-library/notices-of-proposed-amendment/npa-2024-06>
- [29] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*, 2nd ed. Lincoln, Mass: Ganga-Jamuna Press, 2006.

- [30] B. Jaduszliwer and J. Camparo, “Past, present and future of atomic clocks for GNSS,” *GPS Solutions*, vol. 25, pp. 1–13, 2021.
- [31] Interface Control Working Group, “NAVSTAR GPS Space Segment/Navigation User Segment Interfaces, IS-GPS-200, rev. N,” Space Systems Command, Tech. Rep., August 2022.
- [32] E. D. Kaplan and C. Hegarty, *Understanding GPS: principles and applications*, 2nd ed. Artech house, 2006.
- [33] J. Lee, S. Pullen, and P. Enge, “Sigma Overbounding using a position domain method for the local area augmentation of GPS,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 4, pp. 1262–1274, Oct. 2009.
- [34] A. Kleusberg, “Atmospheric models from GPS,” in *GPS for Geodesy*. Springer, 1998, pp. 599–623.
- [35] M. M. Hoque and N. Jakowski, “Estimate of higher order ionospheric errors in GNSS positioning,” *Radio Science*, vol. 43, no. 05, pp. 1–15, 2008.
- [36] M. Z. Jacobson, *Fundamentals of atmospheric modeling*. Cambridge University Press, 1999.
- [37] RTCA Special Committee 159, “RTCA/DO-229D: Minimum operational performance standards for global positioning system/wide area augmentation system airborne equipment,” Radio Technical Commission for Aeronautics (RTCA), Tech. Rep., 2006.
- [38] J. W. Betz, “Fundamentals of satellite-based navigation and timing,” in *Position, Navigation, and Timing Technologies in the 21st Century*, Y. J. Morton, F. van Diggelen, J. J. Spilker Jr, B. W. Parkinson, S. Lo, and G. Gao, Eds. John Wiley & Sons, Ltd, 2020, ch. 2, pp. 43–64.

- [39] M. Smyrnaios, S. Schn, M. Liso, and S. Jin, “Multipath propagation, characterization and modeling in GNSS,” *Geodetic Sciences-Observations, Modeling and Applications*, pp. 99–125, 2013.
- [40] T. Murphy, M. Harris, J. Booth, P. Geren, T. Pankaskie, B. Clark, J. Burns, and T. Urda, “Results from the program for the investigation of airborne multipath errors,” in *Proceedings of the 2005 National Technical Meeting of The Institute of Navigation*, 2005, pp. 153–169.
- [41] Working Group C-ARAIM Technical Subgroup, “Milestone 2 report,” EU-US Cooperation in Satellite Navigation, Tech. Rep., 2015.
- [42] X. Niu, Q. Chen, Q. Zhang, H. Zhang, J. Niu, K. Chen, C. Shi, and J. Liu, “Using Allan variance to analyze the error characteristics of GNSS positioning,” *GPS Solutions*, vol. 18, pp. 231–242, 2014.
- [43] L. Zhao, J. Zhang, L. Li, F. Yang, and X. Liu, “Position-domain non-Gaussian error overbounding for ARAIM,” *Remote Sensing*, vol. 12, no. 12, p. 1992, 2020.
- [44] S. Perea Diaz, “Design of an integrity support message for offline advanced RAIM,” Ph.D. dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, 2019.
- [45] O. Montenbruck, P. Steigenberger, and A. Hauschild, “Broadcast versus precise ephemerides: a multi-GNSS perspective,” *GPS Solutions*, vol. 19, pp. 321–333, 2015.
- [46] T. Walter, J. Blanch, and P. Enge, “Evaluation of signal in space error bounds to support aviation integrity,” *NAVIGATION*, vol. 57, no. 2, pp. 101–113, 2010.
- [47] Y. Wu, X. Liu, W. Liu, J. Ren, Y. Lou, X. Dai, and X. Fang, “Long-term behavior and statistical characterization of BeiDou signal-in-space errors,” *GPS Solutions*, vol. 21, pp. 1907–1922, 2017.

- [48] O. Montenbruck and P. Steigenberger, “GNSS orbit determination and time synchronization,” in *Position, Navigation, and Timing Technologies in the 21st Century*, Y. J. Morton, F. van Diggelen, J. J. Spilker Jr, B. W. Parkinson, S. Lo, and G. Gao, Eds. John Wiley & Sons, Ltd, 2020, ch. 11, pp. 233–258.
- [49] “GPS Constellation Information,” accessed on Sep 2024. [Online]. Available: <https://www.navcen.uscg.gov/gps-constellation>
- [50] “Galileo Constellation Information,” accessed on Sep 2024. [Online]. Available: <https://www.gsc-europa.eu/system-service-status/constellation-information>
- [51] “International GNSS Service, GPS Data Repository,” accessed on May 2024. [Online]. Available: <https://cddis.nasa.gov/archive/gps/data>
- [52] O. Montenbruck, P. Steigenberger, R. Khachikyan, G. Weber, R. Langley, L. Mervart, and U. Hugentobler, “IGS-MGEX: preparing the ground for multi-constellation GNSS science,” *Inside GNSS*, vol. 9, no. 1, pp. 42–49, 2014.
- [53] “International GNSS Service, GNSS Products Repository,” accessed on June 2024. [Online]. Available: <https://cddis.nasa.gov/archive/gnss/products>
- [54] J. Kouba and P. Héroux, “Precise point positioning using igs orbit and clock products,” *GPS Solutions*, vol. 5, pp. 12–28, 2001.
- [55] O. Montenbruck, P. Steigenberger, and A. Hauschild, “Multi-GNSS signal-in-space range error assessment–methodology and results,” *Advances in Space Research*, vol. 61, no. 12, pp. 3020–3038, 2018.
- [56] National Geospatial-Intelligence Agency, “NGA GNSS Division Precise Ephemeris Parameters,” accessed on June 2024. [Online]. Available: <https://earth-info.nga.mil/php/download.php?file=gnss-precise>

- [57] European GNSS Agency, “Galileo metadata,” accessed on July 2024. [Online]. Available: <https://www.gsc-europa.eu/support-to-developers/galileo-satellite-metadata>
- [58] International GNSS Service, “Antenna files,” accessed on July 2024. [Online]. Available: <https://igs.org/wg/antenna/#files>
- [59] O. Montenbruck, R. Schmid, F. Mercier, P. Steigenberger, C. Noll, R. Fatkulin, S. Kogure, and A. S. Ganeshan, “GNSS satellite geometry and attitude models,” *Advances in Space Research*, vol. 56, no. 6, pp. 1015–1029, 2015.
- [60] Department of Defense, *Global Positioning System Standard Positioning Service Performance Standard*, 5th ed., 2020.
- [61] T. Walter, K. Gunning, R. Eric Phelts, and J. Blanch, “Validation of the unfaulted error bounds for ARAIM,” *Navig. J. Inst. Navig.*, vol. 65, no. 1, pp. 117–133, 2018.
- [62] European Union Agency for the Space Programme, *Galileo Open Service Service Definition Document*, 2023.
- [63] B. DeCleene, “Defining pseudorange integrity - overbounding,” in *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, Salt Lake City, UT, Sep. 2000, pp. 1916–1924.
- [64] J. Rife, S. Pullen, B. Pervan, and P. Enge, “Paired overbounding and application to GPS augmentation,” in *PLANS 2004. Position Location and Navigation Symposium (IEEE Cat. No.04CH37556)*, 2004, pp. 439–446.
- [65] J. Blanch, T. Walter, and P. Enge, “Gaussian bounds of sample distributions for integrity analysis,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 4, pp. 1806–1815, Aug. 2019.

- [66] J. Rife, T. Walter, and J. Blanch, “Overbounding SBAS and GBAS error distributions with excess-mass functions,” in *Proceedings of the International Symposium on GNSS/GPS, Sydney, Australia*, Sydney, Australia, 2004.
- [67] C. A. Shively, “A comparison of LAAS error bounding concepts,” in *Proceedings of the 2001 National Technical Meeting of The Institute of Navigation*, Long Beach, CA, Jan. 2001, pp. 501–511.
- [68] J. D. Larson, D. Gebre-Egziabher, and J. H. Rife, “Gaussian-Pareto overbounding of DGNSS pseudoranges from CORS,” *NAVIGATION*, vol. 66, no. 1, pp. 139–150, Jan. 2019.
- [69] S. Coles, “Threshold models,” in *An Introduction to Statistical Modeling of Extreme Values*. London: Springer London, 2001, pp. 74–91. [Online]. Available: https://doi.org/10.1007/978-1-4471-3675-0_4
- [70] J. Pickands III, “Statistical inference using extreme order statistics,” *the Annals of Statistics*, pp. 119–131, 1975.
- [71] J. Blanch, T. Walter, and P. Enge, “Position error bound calculation for GNSS using measurement residuals,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 977–984, Jul. 2008.
- [72] Y. Yun, H. Yun, D. Kim, and C. Kee, “A Gaussian sum filter approach for DGNSS integrity monitoring,” *J. Navig.*, vol. 61, no. 4, pp. 687–703, 2008.
- [73] J. Lee, “LAAS position domain monitor analysis and test results for CAT II/III operations,” in *Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2004)*, Long Beach, CA, Sep. 2004, pp. 2786–2796.
- [74] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, Sep. 1977.

- [75] Z. Gao, K. Fang, Z. Wang, K. Guo, and Y. Liu, “An error overbounding method based on a Gaussian mixture model with uncertainty estimation for a dual-frequency ground-based augmentation system,” *Remote Sensing*, vol. 14, no. 5, p. 1111, 2022.
- [76] J. A. Rice, “Probability,” in *Mathematical Statistics and data analysis*. Thomson Brooks/Cole, 2007, p. 1–26.
- [77] W. C. Horrace, “Moments of the truncated normal distribution,” *Journal of Productivity Analysis*, vol. 43, pp. 133–138, 2015.
- [78] S. Shah and M. Jaiswal, “Estimation of parameters of doubly truncated normal distribution from first four sample moments,” *Annals of the Institute of Statistical Mathematics*, vol. 18, no. 1, pp. 107–111, 1966.
- [79] R. N. Bracewell, *The Fourier Transform and Its Applications*, 3rd ed., ser. McGraw-Hill series in electrical and computer engineering. Boston: McGraw Hill, 2000.
- [80] J. Rife and B. Pervan, “Overbounding revisited: discrete error-distribution modeling for safety-critical GPS navigation,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1537–1551, 2012.
- [81] H. J. Nussbaumer, “The fast Fourier transform,” in *Fast Fourier Transform and Convolution Algorithms*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1982, pp. 80–111. [Online]. Available: https://doi.org/10.1007/978-3-642-81897-4_4
- [82] T. Walter and J. Blanch, “Keynote-Characterization of GNSS clock and ephemeris errors to support ARAIM,” in *Proceedings of the ION 2015 Pacific PNT Meeting*, 2015, pp. 920–931.
- [83] O. Osechas, P. Misra, and J. Rife, “Carrier-phase acceleration RAIM for GNSS satellite clock fault detection,” *NAVIGATION*, vol. 59, no. 3, pp. 221–235, 2012.

- [84] M. Joerger and B. Pervan, “Fault detection and exclusion using solution separation and Chi-squared ARAIM,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 2, pp. 726–742, Apr. 2016.
- [85] R. Wang, Z. Xiong, J. Liu, J. Xu, and L. Shi, “Chi-square and SPRT combined fault detection for multisensor navigation,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 3, pp. 1352–1365, 2016.
- [86] Y. Liu and Y. J. Morton, “Improved automatic detection of GPS satellite oscillator anomaly using a machine learning algorithm,” *NAVIGATION*, vol. 69, no. 1, 2022.
- [87] Z. Gao, S. X. Ding, and C. Cecati, “Real-time fault diagnosis and fault-tolerant control,” *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3752–3756, 2015.
- [88] B. S. Pervan, D. G. Lawrence, C. E. Cohen, and B. W. Parkinson, “Parity space methods for autonomous fault detection and exclusion using GPS carrier phase,” in *Proceedings of Position, Location and Navigation Symposium-PLANS’96*. IEEE, 1996, pp. 649–656.
- [89] M. Joerger and B. Pervan, “Kalman filter-based integrity monitoring against sensor faults,” *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 2, pp. 349–361, 2013.
- [90] J. W. Tukey, “Bias and confidence in not quite large samples,” *Annals of Mathematical Statistics*, vol. 29, p. 614, 1958.
- [91] M. H. Quenouille, “Notes on bias in estimation,” *Biometrika*, vol. 43, no. 3/4, pp. 353–360, 1956.
- [92] M. H. Quenouille, “Problems in plane sampling,” *The Annals of Mathematical Statistics*, vol. 20, no. 3, pp. 355–375, 1949.

- [93] C. Bonferroni, “Teoria statistica delle classi e calcolo delle probabilit ,” *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, vol. 8, pp. 3–62, 1936.
- [94] S. Hutsell, M. Forsyth, and C. B. McFarland, “One-way GPS time transfer: 2002 performance,” in *Proceedings of the 34th Annual Precise Time and Time Interval Systems and Applications Meeting*, 2002, pp. 69–76.
- [95] L. Carlone, A. Censi, and F. 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*. IEEE, 2014, pp. 2667–2674.
- [96] S.-S. Jan, W. Chan, T. Walter, and P. Enge, “MATLAB simulation toolset for SBAS availability analysis,” in *Proceedings of the 14th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2001)*, 2001, pp. 2366–2375.
- [97] U.S. Coast Guard Navigation Center, “GPS Almanac File at Week 2243,” accessed on Oct 2024. [Online]. Available: <https://www.navcen.uscg.gov/sites/default/files/gps/almanac/2023/Yuma/001.alm>
- [98] European GNSS Service Center, “Galileo Almanac File at Week 2243,” accessed on Oct 2024. [Online]. Available: <https://www.gsc-europa.eu/sites/default/files/sites/all/files/2023-01-03.xml>
- [99] L. Yang, N. L. Knight, Y. Li, and C. Rizos, “Optimal fault detection and exclusion applied in GNSS positioning,” *The Journal of Navigation*, vol. 66, no. 5, pp. 683–700, 2013.
- [100] EU US Working Group C, “WG-C Advanced RAIM Technical Subgroup Reference Airborne Algorithm Description Document,” accessed on April

2024. [Online]. Available: https://web.stanford.edu/group/scpnt/gpslab/website_files/maast/ARAIM_TSG_Reference_ADD_v3.0.pdf
- [101] S. Holm, “A simple sequentially rejective multiple test procedure,” *Scandinavian Journal of Statistics*, pp. 65–70, 1979.
- [102] G. A. McGraw, T. Murphy, M. Brenner, S. Pullen, and A. Van Dierendonck, “Development of the LAAS accuracy models,” in *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, 2000, pp. 1212–1223.
- [103] J. L. W. V. Jensen, “Sur les fonctions convexes et les inégalités entre les valeurs moyennes,” *Acta Math.*, vol. 30, no. 1, pp. 175–193, 1906.
- [104] J. P. Mills, “Table of the ratio: area to bounding ordinate, for any portion of normal curve,” *Biometrika*, pp. 395–400, Nov. 1926.
- [105] W. Greene, *Econometric Analysis*, 4th ed. Upper Saddle River, NJ: Prentice Hall Internat, 2000.
- [106] A. C. Aitken, “On least squares and linear combination of observations,” *Proceedings of the Royal Society of Edinburgh*, vol. 55, pp. 42–48, 1936.