DGK Ausschuss Geodäsie (DGK)
der Bayerischen Akademie der Wissenschaften

Reihe C | Dissertationen | Heft Nr. 965

Jingyao Su

Towards interval-based autonomous integrity monitoring: Error bounding and uncertainty propagation

München 2025

Verlag der Bayerischen Akademie der Wissenschaften, München

ISSN 0065-5325 ISBN 978-3-7696-5377-9

DGK Ausschuss Geodäsie (DGK)
der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 965

Towards interval-based autonomous integrity monitoring: Error bounding and uncertainty propagation

Von der Fakultät für Bauingenieurwesen und Geodäsie
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades
Doktor-Ingenieur (Dr.-Ing.)
genehmigte Dissertation

von

Jingyao Su, M. Sc.

München 2025

Verlag der Bayerischen Akademie der Wissenschaften, München

ISSN 0065-5325 ISBN 978-3-7696-5377-9

Adresse des Ausschusses Geodäsie (DGK)
der Bayerischen Akademie der Wissenschaften:

Ausschuss Geodäsie (DGK) der Bayerischen Akademie der Wissenschaften
Alfons-Goppel-Straße 11 ● D – 80 539 München
Telefon +49 – 89 – 23 031 1113 ● Telefax +49 – 89 – 23 031 - 1283 / - 1100
e-mail post@dgk.badw.de ● http://www.dgk.badw.de

Prüfungskommission:

Vorsitzender:	Prof. Dr. Matthias Weigelt
Referent:	Prof. Dr.-Ing. Steffen Schön
Korreferenten:	Prof. Dr. Mathieu Joerger
	apl. Prof. Dr. techn. Franz Rottensteiner

Tag der mündlichen Prüfung: 26.03.2025

:

Diese Dissertation ist auf dem Server des Ausschusses Geodäsie (DGK)
der Bayerischen Akademie der Wissenschaften, München unter <http://dgk.badw.de/>
sowie unter Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik
der Leibniz Universität Hannover (ISSN 0174-1454), Nr. 410,
<https://doi.org/10.15488/19619>, Hannover 2025, elektronisch publiziert

ISSN 0065-5325 ISBN 978-3-7696-5377-9

Abstract

Satellite navigation can provide essential positioning, navigation and timing (PNT) information to a broad range of users. With the development of the Global Navigation Satellite System (GNSS), various applications have emerged and grown with substantial economic impact over the past decades. For safety-critical GNSS applications, estimation errors must be reliably quantified and safely bounded. This requirement is crucial for ensuring navigation integrity, which was originally formulated for aviation navigation. It concerns the trust that can be placed in a navigation solution under rare-event conditions and, hence, differs from the commonly expressed user demand in accuracy, focusing on the trustiness and reliability of the navigation system. Conventional solutions have been concentrating on stochastic approaches, relying on distributional assumptions for the observation errors before they are propagated through state estimation. However, the exact error distribution is either unknown or hardly validated, and the remaining systematics persist in the GNSS measurements after applying correction methods. In this regard, purely stochastic modeling of all error sources will not be adequate, necessitating the exploration of alternative approaches.

Interval is a promising alternative representation of uncertainty. It provides deterministic bounds that indicate the possible variation of errors and, hence, is feasible to represent the uncertainty due to remaining systematic effects. Grounded on the interval-described uncertainty modeling, an innovative integrity monitoring framework is developed in this dissertation, providing an alternative approach to classical stochastic methods such as (Advanced) Receiver Autonomous Integrity Monitoring (RAIM and ARAIM). Critical integrity-focused questions include: (i) how representative the navigation solutions are and (ii) how their uncertainty can be safely modeled to yield integrity assurance.

To address these questions, practical methods of determining interval bounds for various GNSS error sources are developed and validated through experiments. Building on the interval bounds, the state estimation problem is investigated, with an emphasis on error bounding. Two novel point estimators are proposed by exploring the set-described uncertainty models, showing advantageous error bounding performance. The set-based fault detector, intended to be integrated into the new integrity monitoring architecture, is discussed and assessed in comparison with classical methods. By utilizing the multiple hypotheses framework that is also adopted by ARAIM, the developed approach can handle multiple simultaneous faults, protecting the navigation system from loss of integrity. Its effectiveness is demonstrated by various evaluation strategies and compared to state-of-art methods, including analytical and Monte Carlo assessments, as well as performance analysis with real-world experiments. Additionally, this dissertation proposes to improve the baseline ARAIM algorithms by implementing the interval extension of the least-squares estimator.

Keywords: GNSS, integrity, uncertainty, interval mathematics, set theory

Zusammenfassung

Satellitennavigation kann einer breiten Palette von Benutzern wichtige positioning, navigation and timing (PNT)-Informationen liefern. Mit der Entwicklung des Global Navigation Satellite System (GNSS) sind in den letzten Jahrzehnten zahlreiche Anwendungen entstanden und gewachsen, die erhebliche wirtschaftliche Auswirkungen hatten. Für sicherheitskritische GNSS-Anwendungen müssen Schätzfehler zuverlässig quantifiziert und sicher begrenzt werden. Diese Anforderung ist entscheidend für die Gewährleistung der Navigationsintegrität und wurde ursprünglich für die Luftfahrtnavigation formuliert. Sie betrifft das Vertrauen, das man unter seltenen Ereignisbedingungen in eine Navigationslösung setzen kann, und unterscheidet sich daher von der allgemein geäußerten Genauigkeitsanforderung der Benutzer, die sich auf die Vertrauenswürdigkeit und Zuverlässigkeit des Navigationssystems konzentriert. Herkömmliche Lösungen konzentrierten sich auf stochastische Ansätze und verließen sich auf Verteilungsannahmen für die Beobachtungsfehler, bevor diese durch die Zustandsschätzung weitergegeben wurden. Die genaue Fehlerverteilung ist jedoch entweder unbekannt oder kaum validiert, und die verbleibende Systematik bleibt nach Anwendung von Korrekturmethoden in den GNSS-Messungen bestehen. In dieser Hinsicht ist eine rein stochastische Modellierung aller Fehlerquellen nicht ausreichend, sodass alternative Ansätze untersucht werden müssen.

Intervalle sind eine vielversprechende alternative Darstellung der Unsicherheit. Sie bieten deterministische Grenzen, die die mögliche Fehlervariation anzeigen, und sind daher geeignet, die Unsicherheit aufgrund verbleibender systematischer Effekte darzustellen. Basierend auf der intervallbeschriebenen Unsicherheitsmodellierung wird in dieser Dissertation ein innovatives Integritätsüberwachungsframework entwickelt, das einen alternativen Ansatz zu klassischen stochastischen Methoden wie (Advanced) Receiver Autonomous Integrity Monitoring (RAIM und ARAIM) bietet. Kritische integritätsbezogene Fragen umfassen: (i) wie repräsentativ die Navigationslösungen sind und (ii) wie ihre Unsicherheit sicher modelliert werden kann, um Integritätssicherheit zu gewährleisten.

Um diese Fragen zu beantworten, werden praktische Methoden zur Bestimmung von Intervallgrenzen für verschiedene GNSS-Fehlerquellen entwickelt und durch Experimente validiert. Aufbauend auf den Intervallgrenzen wird das Problem der Zustandsschätzung untersucht, wobei der Schwerpunkt auf der Fehlerbegrenzung liegt. Durch die Untersuchung der mengenbeschriebenen Unsicherheitsmodelle werden zwei neuartige Punktschätzer vorgeschlagen, die eine vorteilhafte Fehlerbegrenzungsleistung aufweisen. Der mengenbasierte Fehlerdetektor, der in die neue Integritätsüberwachungsarchitektur integriert werden soll, wird im Vergleich mit klassischen Methoden diskutiert und bewertet. Durch die Verwendung des Mehrfachhypothesenrahmens, der auch von ARAIM übernommen wird, kann der entwickelte Ansatz mehrere gleichzeitige Fehler verarbeiten und das Navigationssystem vor Integritätsverlust schützen. Seine Wirksamkeit wird durch verschiedene Bewertungsstrategien nachgewiesen und mit modernsten Methoden verglichen, darunter analytische und Monte Carlo-Bewertungen sowie Leistungsanalysen mit realen Experimenten. Darüber hinaus schlägt diese Dissertation vor, die grundlegenden ARAIM-Algorithmen durch Implementierung der Intervallerweiterung des least-squares-Schätzers zu verbessern.

Schlüsselwörter: GNSS, Integrität, Unsicherheit, Intervallmathematik, Mengenlehre

Contents

List of Figures

List of Tables

Acronyms

ABAS	Aircraft-based Augmentation System.
ACF	Autocorrelation Function.
ADAS	Advanced Driver Assistance Systems.
AL	Alert Limit.
ARAIM	Advanced Receiver Autonomous Integrity Monitoring.
BDGIM	BeiDou Global broadcast Ionospheric delay correction Model.
BDS	BeiDou Navigation Satellite System.
BLUE	best linear unbiased estimation.
BMR	bound-minus-residual.
BOC	binary offset carrier.
BPSK	binary phase shift keying.
BRDC	broadcast ephemerids products.
C/A	coarse/acquisition.
CAS	Chinese Academy of Sciences.
CBOC	composite binary offset carrier.
CDF	Cumulative Density Function.
CLT	Central Limit Theorem.
CNMP	receiver node noise and multipath.
CODE	the Center for Orbit Determination in Europe.
CR	continuity risk.
CSP	Constellation Service Provider.
DIA	detection, identification and adaptation.
DLL	Delay Locked Loop.
DOP	dilution of precision.
DOY	Day-Of-Year.
DWD	Deutscher Wetterdienst.
ECDF	Empirical Cumulative Density Function.
EDM	elevation-dependent model.
EEM	elementary error model.
EOPP	Earth Orientation Prediction Parameter.
EPD	extra path delay.
EUS	estimate uncertainty set.
EUSPA	European Union Agency for the Space Programme.
FA	false alert.

FAA Federal Aviation Administration.

FD fault detection.

FDE fault detection and exclusion.

FE fault exclusion.

FSS feasible solution set.

GBAS Ground-based Augmentation System.

GEO Geostationary Earth Orbit.

GIM global ionosphere (TEC) maps.

GLONASS Global Orbiting Navigation System/Global'naya Navigatsionnaya Sputnikovaya Sistema.

GNSS Global Navigation Satellite System.

GPS Global Positioning System.

GPT Global Pressure and Temperature model.

GUM Guide to the expression of uncertainty in measurement.

HI Hazardous Information.

HMI Hazardous Misleading Information.

HPE Horizontal Position Error.

HPL Horizontal Protection Level.

HUL Horizontal Uncertainty Level.

ICAO International Civil Aviation Organization.

ICLS inequality-constrained least-squares.

IF Ionosphere-Free Linear Combination.

IfE Institut für Erdmessung.

IGS International GNSS Service.

iid independent identically distributed.

IMU Inertial Measurement Unit.

inid independent nonidentically distributed.

INS Inertial Navigation System.

IONEX the IONosphere Map EXchange Format.

IPP ionospheric pierce points.

IR integrity risk.

ISM Integrity Support Message.

ISO International Standard Organization.

ISP Integrity Support Parameter.

IURE Instantaneous User Range Error.

LAD least-absolute-deviation.

LC Linear Combination.

LOC loss of continuity.

LOI loss of integrity.

LOS Line-of-Sight.

LSQ least-squares.

MBOC multiplexed binary offset carrier.

MC Monte Carlo.

MD missed detection.

MDB Minimum Detectable Bias.

MEE multipath error envelope.

MEO Medium Earth Orbit.

MHSS Multiple Hypotheses Solution Separation.

MI Misleading Information.

MJD Modified Julian Date.

MOPS Minimum Operational Performance Standard.

MUS measurement uncertainty set.

ND No Detection.

NLOS Non-Line-of-Sight.

NLS Non-Least-Squares.

NO Normal Operation.

NP No Protection.

NSP Navigation Systems Panel.

OCS Operational Control Segment.

OMC Observed-Minus-Computed.

OS Open Service.

PDF Probability Density Function.

PE Position Error.

PF Position Failure.

PL Protection Level.

PNT positioning, navigation and timing.

PRN Pseudo-Random Noise.

PSD Power Spectral Density.

RAIM Receiver Autonomous Integrity Monitoring.

RB residual-based.

RF Radio Frequency.

RINEX Receiver Independent Exchange.

RMS root mean squared.

RT ray tracing.

RTCA Radio Technical Commission for Aeronautics.

SBAS Satellite-based Augmentation System.

SCLS set-constrained least-squares.

SDM signal strength-dependent model.

SISA Signal-In-Space Accuracy.

SISE Signal-In-Space Error.

SISRE Signal-In-Space Range Error.

SIVIA Set Inversion via Interval Analysis.

SPD Slant Path Delay.

SPP Single Point Positioning.

SS solution separation.

STEC slant total electron content.

SU System Unavailable.

TEC total electron content.

TECU total electron content unit.

TTA Time-to-Alert.

UCP User Consultation Platform.

UIRE User Ionospheric Range Error.

UIVE User Ionospheric Vertical Error.

URA User Range Accuracy.

URE User Range Error.

VCM Variance-Covariance Matrix.

VIM International Vocabulary of Metrology.

VPL Vertical Protection Level.

VTEC vertical total electron content.

WAAS Wide Area Augmentation System.

WCF worst-case fault.

ZHD Zenith Hydrostatic Delay.

ZPD Zenith Path Delay.

ZTD Zenith Total Delay.

ZWD Zenith Wet Delay.

Notations and Nomenclatures

Constants and Physical Parameters

a	signal amplitude.
ϕ	azimuth angle.
ψ	Earth-centered zenith angle.
θ	elevation angles.
c	speed of light in vacuum, $c = 299\,792\,458\,\mathrm{m/s}$.
$\alpha_i,\ \beta_i$	model constants, e.g., in the Saastamoinen model and GPS Klobuchar model.
d_c	GNSS receiver correlator spacing in the unit of PRN chips.
f	signal frequency, e.g., for GPS L1/L2/L5: $f_1 = 1575.42\,\mathrm{MHz}, f_2 = 1227.6\,\mathrm{MHz}, f_5 = 1176.45\,\mathrm{MHz}$.
H	height.
Φ	latitude. Variants include: subionospheric latitude Φ_I, geomagnetic latitude Φ_m.
Λ	longitude. A variant is the subionospheric longitude Λ_I.
ω	signal-to-noise (M/D) ratio.
e	partial water vapor pressure.
p	air pressure.
T	temperature.
ϑ	relative phase.
Γ	autocorrelation sidelobe level for PRN codes. For example, Γ_{max} and Γ_{min} denote the maximum and minimum sidelobes.
$s_i(t)$	i-th signal component at epoch t.
T_{BOC}	chip duration of the BOC signal.
$T_{C/A}$	chip duration of the C/A signal.
t	time.
z	apparent (radio) zenith distance of the satellite.
Z	true (radio) zenith distance of the satellite.

GNSS Measurements and Error Sources

$\delta t_r(t)$	receiver clock offset at epoch t.
$\delta t^k(t)$	satellite clock offset at epoch t.
τ_r^k	apparent signal travel time from satellite k to receiver r.
$M(\theta)$	elevation-dependent slant factor, or mapping function, to map the atmospheric zenith delays to slant delays.
D	extra path delay (EPD).
ϵ_{clk}	instantaneous satellite clock error.

I_r^k	ionospheric delay.
$n_e(l)$	electron density along the signal propagation path l.
$\epsilon_{IURE,r}^k$	Instantaneous User Range Error (IURE) between satellite k and receiver r.
$\epsilon_{MP,r}^k$	pseudorange multipath error.
$\epsilon_{NLOS,r}^k$	pseudorange measurement error due to Non-Line-of-Sight (NLOS) extra path delay (EPD).
ϵ_{noise}	noise and other remaining error terms in the GNSS observation equations.
ϵ_{orb}^k	instantaneous satellite ephemeris error vector for satellite k.
T_r^k	tropospheric delay.
$n_T(l)$	tropospheric refractivity along the signal propagation path l.
$\mathbf{l}_r^k$	unit vector of the LOS direction between satellite k and receiver r.
φ	carrier phase measurement.
p	pseudorange measurement. Particularly, $p_r^k(t)$ denotes the measurement between satellite k and receiver r at epoch t.
ζ	bound-minus-residual (BMR) values, e.g., ζ_I denotes the ionospheric BMR, ζ_T the tropospheric BMR, and ζ_{MP} the multipath BMR.
δ	residual values after applying correction methods, such as δ_I denotes the ionospheric residuals and δ_T the tropospheric residuals.
$\rho_r^k(t)$	geometric distance between satellite k and receiver r at epoch t.
$\mathbf{p}^k = [x^k, y^k, z^k]^T$	position vector of satellite k.
$\mathbf{p}_r = [x_r, y_r, z_r]^T$	position vector of receiver r. A variant is $\mathbf{p}_{r,0}$ for initial position estimate of the receiver.
$\Delta\mathbf{p}_r = \mathbf{p}_r - \mathbf{p}_{r,0}$	vector of receiver's coordinate increment with respect to its initial position estimate.

Integrity Monitoring Parameters

ℓ	Alert Limit (AL).
b_{nom}	nominal bias.
k	multipier for computing Protection Level (PL), including k_H and k_V for horizontal and vertical PL in RB RAIM, respectively.
$P(\cdot)$	probability of an event. Examples include: P(FA) for false alert (FA) probability, P(MD) for missed detection (MD) probability, P(NP) for No Protection (NP) probability, P(HI) for Hazardous Information (HI) probability, P(PF) for Position Failure (PF) probability, P(HMI) for Hazardous Misleading Information (HMI) probability, etc..
P_{const}	constellation fault probability.
P_Δ	bounding probability associated with the interval error bounds $[-\Delta, \Delta]$.
P_{event}	probability of an independent fault event, which can be a single satellite failure or a constellation failure.
P_{NM}	propability that more than r simultaneous satellite failures occur but are not monitored..
P_{sat}	satellite fault probability.
q_{RB}	residual-based (RB) test statistic.
q_{SS}	solution separation (SS) test statistic.
C_{REQ}	continuity risk (CR) requirement.
I_{REQ}	integrity risk (IR) requirement.

$\sigma_{\mathbf{bias}}$	standard deviation used to calculate Horizontal Protection Level (HPL) in RB RAIM.
SLOPE	slope term. Variants include $\text{HLOPSE}_{\max}$ and $\text{VSLOPE}_{\max}$ for the maximum horizontal and vertical slopes, respevtively.
Δ_{SS}	solution separation (SS).
T_{RB}	residual-based (RB) test threshold.
T_{SS}	solution separation (SS) test threshold.

Interval and Set Representations

$[y] = [\underline{y}, \overline{y}]$	an interval and its lower bound $\underline{y}$ and upper bound $\overline{y}$.
Δ	interval radius, e.g., Δ_d for the detection interval and Δ_s for the solution interval.
$\mathbf{G}_{\mathcal{Z}}$	generator matrix of a zonotope with its i-th column vector $\mathbf{g}_{\mathcal{Z}}^{(i)}$.
$\mathcal{C}_C$	constrained confidence region for the set-constrained least-squares (SCLS).
$\mathcal{C}_E$	extended confidence region for the least-squares estimator.
$\mathcal{E}$	ellipsoid, e.g., $\mathcal{E}_{1-\alpha}$ represents the $1-\alpha$ confidence ellipsoid.
$\mathcal{P}$	polytope, e.g., $\mathcal{P}_s$ represents the solution polytope and $\mathcal{P}_D$ the detection polytope.
$\mathcal{S}$	feasible solution set (FSS).
$\mathcal{X}$	a general set.
$\mathcal{Z}$	zonotope. Variants include: $\mathcal{Z}_s$ denotes the confidence zonotope, i.e., measurement uncertainty set (MUS), for the set estimator, and $\mathcal{Z}_D$ denotes the zonotope that describes the uncertainty due to remaining systematic error for the least-squares estimator.

Operators and Relations

$\in$	set belong.
$\notin$	set not belong.
$\text{col}(\cdot)$	the column vector of a matrix.
$\text{conv}(\cdot)$	the convex hull of a set.
$\text{diag}(\cdot)$	define a diagonal matrix with the diagonal components of the given matrix.
$:=$	is defined as.
$\text{wid}(\cdot)$	the width of an interval.
$\wedge$ **or** $\bigwedge$	logical *and*.
$\vee$ **or** $\bigvee$	logical *or*.
$\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$	maximum output of the function $f(\mathbf{x})$ for all possible input $\mathbf{x} \in \mathcal{X}$.
$\text{meas}(\cdot)$	an (arbitrary) scalar measure of a set.
$\text{mid}(\cdot)$	the midpoint of an interval. Alternatively, the midpoint of $[y]$ can be denoted by y_m.
$\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$	minimum output of the function $f(\mathbf{x})$ for all possible input $\mathbf{x} \in \mathcal{X}$.
$\lvert\cdot\rvert_p$	p-norm operation, specially, $\lvert\cdot\rvert$ for 1-norm, and $\lVert\cdot\rVert$ for 2-norm.
$\text{cen}(\cdot)$	(Chebyshev) center of a set .
$\text{diam}(\cdot)$	diameter of a set .
$\cap$ **or** $\bigcap$	set intersection.

$\oplus$	Minkowski sum.
$\mathrm{rad}(\cdot)$	the radius of a set or an interval.
$\cup$ **or** $\bigcup$	set union.
$\subseteq$	subset.
$\nsubseteq$	not a subset.
$\subsetneq$ **or** $\subset$	proper subset, i.e., subset but not equal to.
$\leq$ **and** $\geq$	in addition to classical scalar inequalities, this thesis also adopts them for vector inequalities, or component-wise inequality in R^n: $\mathbf{u} \leq \mathbf{v}$ means $\forall u_i \leq v_i$ for $i = 1, ..., n$, and $\mathbf{u} \geq \mathbf{v}$ means $\forall u_i \geq v_i$ for $i = 1, ..., n$.
$\nleq$ **and** $\ngeq$	denote vector inequalities or component-wise inequality in this thesis: $\mathbf{u} \nleq \mathbf{v}$ means $\exists u_i > v_i$ for $i = 1, ..., n$, and $\mathbf{u} \ngeq \mathbf{v}$ means $\exists u_i < v_i$ for $i = 1, ..., n$.

Probability Distribution and Uncertainty Modeling

$f_N(\cdot)$	Probability Density Function (PDF) of a normal distribution $\mathcal{N}(0, 1)$.
$\chi^2(\nu, \lambda)$	non-central chi-square distribution, with the degree of freedom ν, and non-centrality λ.
$\mathcal{N}(\mu, \sigma^2)$	a Gaussian distribution with mean μ and variance σ^2.
$\mathcal{N}(\mu, \mathbf{\Sigma})$	a multivariate Gaussian distribution with mean vector μ and covariance matrix $\mathbf{\Sigma}$.
$F_N(\cdot)$	Cumulative Density Function (CDF) of a normal distribution $\mathcal{N}(0, 1)$.
$Q(\cdot)$	tail probability distribution function: $Q\{\} = 1 - F_N\{\}$.
$Q^{-1}(\cdot)$	inverse of the tail probability distribution function, i.e., $Q^{-1}(P)$ is the $(1 - P)$ quantile of a normal distribution.
$t(\nu)$	student's t distribution, with the degree of freedom ν.
$\mathcal{U}_{[\underline{u}, \overline{u}]}$	uniform distribution, with the lower bound $\underline{u}$ and upper bound $\overline{u}$.
$\mathbf{s}$	systematic error term (in contrast to random errors).
$\mathbf{v}$	random error term (in contrast to systematic errors).
$\mathbf{d}$	influence factor vector.
$\mathbf{F}$	matrix of sensitivity coefficients.
$E\{\cdot\}$	mathematical expectation of a random variable, or multi-variables.
$D\{\cdot\}$	dispersion of a random variable, or multi-variables.

State Estimation and Hypothesis Testing

$\mathbf{A}$	design matrix.
$\mathbf{B}$	$\mathbf{B} = [\mathbf{A}; -\mathbf{A}]$ used to form the system of inequalities.
$\mathbf{H}_i$	a $n \times m$ matrix used to extract the fault-free design matrix $\mathbf{A}_i$ under hypothesis H_i, assuming that n_i out of n measurements are faulty, i,e., $\mathbf{A}_i = \mathbf{H}_i^T \mathbf{A}$, $\mathbf{H}_i \coloneqq \begin{bmatrix} \mathbf{I}_{n_i} \\ \mathbf{0}_{(n-n_i)\times n_i} \end{bmatrix}$.
$\mathbf{H}_i^*$	a $n \times m$ matrix used to extract the faulty design matrix $\mathbf{A}_i^*$ under hypothesis H_i, assuming that n_i out of n measurements are faulty, i,e., $\mathbf{A}_i^* = \mathbf{H}_i^{*T} \mathbf{A}$, $\mathbf{H}_i^* \coloneqq \begin{bmatrix} \mathbf{0}_{n_i\times(n-n_i)} \\ \mathbf{I}_{n-n_i} \end{bmatrix}$.
$\mathbf{K}$	transformation matrix of the least-squares estimator.
$\mathbf{P}$	weight matrix.

$\mathbf{Q}$	parity matrix.
$\mathbf{S}$	slope matrix.
$\mathbf{\Sigma}$	covariance matrix, e.g, $\mathbf{\Sigma_{yy}}$ for measurements, $\mathbf{\Sigma_{\hat{x}\hat{x}}}$ for the state estimate.
α^T	vector that is used to extract states of interest, e.g., $\alpha^T = [0, 0, 1, 0]$ for the vertical state using GPS-only signals.
$\mathbf{b}$	vector of upper bounds for the set estimator $\mathbf{b} = [\mathbf{y} - \underline{\mathbf{e}}; -\mathbf{y} + \overline{\mathbf{e}}]$.
$\mathbf{e}$	measurement error vector. Corresponding interval bounds are expressed as $[\mathbf{e}] = [\underline{\mathbf{e}}, \overline{\mathbf{e}}]$.
ε	estimation error vector.
$\mathbf{f}$	unmodeled or unbounded fault vector.
$\mathbf{f_{worst}}$	worst-case fault vector that maximizes the integrity risk.
$\mathbf{k}_{(m)}^T$	row vector of $\mathbf{K}$ containing the least-squares coefficients for the m-th state.
$\mathbf{p}$	parity vector.
$\mathbf{r}$	the least-squares residual vector. A variant is $\tilde{\mathbf{r}}$ for normalized residuals.
$\mathbf{u}$	unit vector of fault line.
$\mathbf{x}$	state vector. A variant is $\hat{\mathbf{x}}$ representing the state estimate.
$\mathbf{y}$	measurement vector. Variants include: $\hat{\mathbf{y}}$ for estimate of measurements and $\tilde{\mathbf{y}}$ for normalized measurements.
H_0/H_A	null and alternative hypotheses in binary hypothesis testing.
H_i	i-th hypothesis in the multiple hypotheses framework for $i = 0, ..., h$.

1 Introduction

Background and motivation. *Satellite navigation* is an integral part of modern life, providing essential positioning, navigation and timing (PNT) information and driving the growth of various applications with substantial economic impact. Since the 1970s, the development of *Global Navigation Satellite System (GNSS)* has continued to advance, leading to the operation of four major GNSS constellations today: the United States' Global Positioning System (GPS), Russia's Global Orbiting Navigation System/Global'naya Navigatsionnaya Sputnikovaya Sistema (GLONASS), the European Galileo system, and China's Beidou System. Each system continues to evolve, delivering increasing capabilities and robustness to meet the growing demands across various sectors.

Uncertainty is a core concern in many scientific and engineering disciplines. In satellite navigation, it is particularly critical for navigation algorithms to incorporate proper models of uncertainty during data processing and to reliably evaluate the uncertainty of navigation solutions. Four key metrics – *accuracy*, *integrity, continuity, and availability* – quantify the navigation performance. These metrics, initially established for aviation, apply to a broad range of navigation scenarios. While high accuracy is a commonly expressed user demand, other metrics, especially integrity, are of vital importance for safety-critical applications such as civil aviation and autonomous driving. Crucial questions for integrity monitoring approaches in these fields include: (i) how representative the navigation solutions are and (ii) how their uncertainty can be safely modeled to yield integrity assurance (which may need to be acceptably conservative). These concerns drive the development of *error bounding* concepts, applicable not only in the observation domain but also in propagation through state estimation.

Conventionally, GNSS error bounding and uncertainty propagation have been mostly handled using stochastic methods. However, in many practical scenarios, the exact error distribution is unknown, and remaining systematics may persist (Schön, 2003). In this regard, purely stochastic modeling of all error sources will not be adequate, necessitating the exploration of alternative approaches. The interval-based approach, also known as set-membership or unknown-but-bounded approach, offers a promising alternative. By leveraging interval mathematics and set theory, this approach describes uncertainties and establishes error bounds as intervals in the observation domain, relaxing the reliance on assumptions about probability distributions. Uncertainties are linearly mapped to the state domain, allowing for uncertainty propagation in a deterministic manner, which makes it suitable for GNSS navigation tasks.

Despite extensive research on interval-based methods in fields such as automatic control and engineering reliability, their application in GNSS has not received much attention. The potential benefits to GNSS navigation, particularly in terms of integrity performance, have yet to be explored. This dissertation seeks to fill this gap by developing a comprehensive interval-

based framework for GNSS integrity monitoring, with a specific focus on urban navigation environments. To achieve this goal, several challenges should be tackled:

1. **Appropriate uncertainty modeling for GNSS observations.** Major error sources must be thoroughly identified and evaluated using appropriate uncertainty models. For instance, it is common practice to assume that the remaining errors after applying various correction methods follow Gaussian distributions. In existing integrity methods such as the (Advanced) Receiver Autonomous Integrity Monitoring (RAIM/ARAIM), Gaussian distributions with inflated standard deviations are typically leveraged to account for rare fault events (RTCA/DO-229D, 2006; WG-C ARAIM TSG, 2019). However, whether this approach still ensures sufficient performance in the presence of systematic errors or non-Gaussian behaviors is questionable, and it may even be unrealistic to validate.

2. **State estimation and corresponding error bounding.** An estimator refers to a function that solves the parameter/state estimation problem such as GNSS navigation, where uncertainties transfer from the observation domain to the parameter/state domain (Koch, 1999). In order to derive error bounds that reflect the factual estimation uncertainty, it is important to understand uncertainty propagation through state estimators. This depends on both the properties of dedicated estimators and developed observation error models. Integrity-focused estimators can be investigated, which is not necessarily optimal in a general sense, e.g., the least-squares estimator can be modified to improve RAIM functionalities (Joerger et al., 2012).

3. **Fault detection and exclusion (FDE) methods.** Robust FDE functions are essential for integrity monitoring. This requires not only strong detection capabilities but also a characterization of the detector in order to ensure integrity in the presence of undetected and unbounded faults. Hypothesis testing, based on self-consistency tests, provides a rigorous framework for navigation fault detection with only GNSS observations. However, the emergence of multi-GNSS systems in complex measurement environments places new challenges, particularly in adapting traditional hypothesis testing to handle multiple simultaneous faults and accommodate novel uncertainty models.

4. **Integrity monitoring algorithms incorporating established uncertainty models.** Integrity monitoring algorithms, tailored to application-specific requirements, should be capable of warning users (e.g., pilot, driver, automated system, etc.) timely when the system is found unsafe to use (i.e., an "alert"). If an alert is not issued, the system must maintain safety, allowing users to compute a position domain error bound to determine whether to accept navigation solutions (Pullen and Joerger, 2021). This requires defining a set of Integrity Support Parameters (ISP) and developing algorithms suited to established uncertainty models, state estimators, FDE methods, and the intended urban navigation environments.

These challenges are addressed in subsequent chapters, where relevant scientific publications will be thoroughly reviewed. The thesis contributions and outline sections provide guidance for navigating the content.

Scientific context. As implied by the title, this dissertation focuses on exploiting interval mathematics for uncertainty modeling and propagation and, subsequently, addressing safety concerns stemming from GNSS integrity requirements across various application scenarios. The work is grounded in the intersection of *interval mathematics*, *set theory*, *adjustment theory*, and *probability theory*. We state in the following how these topics interact within the scope of this thesis:

- **Interval mathematics and set theory.** Interval mathematics, integrated with set-theoretic operations, provides a rigorous framework for computing interval values with lower bounds and upper bounds. The intervals are considered to contain the exact answer to various mathematical problems with absolute certainty, representing bounds on accumulated rounding errors, approximation errors, and propagated uncertainties through measurement models or computational processes (Moore et al., 2009). In general, they are powerful mathematical tools in the course of this study.
- **Adjustment theory.** In many geodetic applications, the unknown parameters to be determined bear a linear (or linearized) relationship to a set of empirical data. The adjustment theory addresses the parameter estimation problem from redundant observational data, ensuring consistency in the presence of errors. This process typically involves stochastic modeling (e.g., the Gauss-Markov model), estimation theory, and hypothesis testing, formulating the fundamental problems in GNSS positioning.
- **Probability theory.** Probability theory provides the mathematical foundation for describing and analyzing random events, which is essential for statistical modeling of measurement errors in GNSS. It underpins traditional approaches to navigation integrity monitoring and is also explored for the developed interval-based approach in this work.

Thesis contributions. In response to the challenges presented, five main contributions of this thesis are outlined below:

- **Development of distribution-free uncertainty models for GNSS pseudorange measurements.** Three major error sources occurring in urban environments are investigated:
  - **Residual tropospheric errors.** Interval uncertainty for the residual tropospheric error is evaluated using sensitivity analysis via interval arithmetic. The Saastamoinen model is studied with the input of meteorological parameters from three models, including the International Standard Organization (ISO) standard atmosphere, the model suggested by Minimum Operational Performance Standard (MOPS), and the Global Pressure and Temperature model (GPT2w).
  - **Residual ionospheric errors.** Interval uncertainty for the residual tropospheric error is evaluated using sensitivity analysis via interval arithmetic. The GPS Klobuchar model is studied with the input of eight broadcast coefficients obtained from the GPS navigation messages. The International GNSS Service (IGS) global ionosphere (TEC) maps (GIM) products are used as references for interval determination.
  - **Multipath errors.** A multipath error envelope (MEE)-based method is proposed to model pseudorange multipath errors for binary phase shift keying (BPSK) and binary offset carrier (BOC) signals. The performance is evaluated and validated using experimental data from GPS, GLONASS, and Galileo observations.
- **Rigorous derivation and evaluation for set-based state estimation.** State estimation in GNSS applications may be approached through the classical least-squares estimator or the set estimator based on constraint satisfaction. For both approaches, the estimation error must be sufficiently bounded for safety reasons. This dissertation revisits error bounding for both estimators and develops two novel point estimators:
  - The **optimal set-based central estimator**, which is optimal under a deterministic optimality concept and provides minimal estimation error bounds.

  - The **set-constrained least-squares estimator**, which integrates set constraints into the classical least-squares framework.

  These estimators are compared through detailed analyses of their error bounding characteristics and performance under various uncertainty models.

- **Exploitation of set-based fault detection.** Unlike traditional fault detection methods that rely on statistical hypothesis testing, the set-based detector operates through set operations, alleviating the need for scalar threshold computations. The mathematical behavior of the set-based detector is analyzed in a simplified scalar estimation scenario using order statistics and compared to classical RAIM detectors, including RB and SS methods. Monte Carlo (MC) evaluations are conducted to demonstrate the performance of the set-based detector in more complex two-dimensional scenarios. An innovative weighting scheme is also proposed to enhance fault detection capabilities.
- **Development of interval-based integrity monitoring architecture against multiple faults.** Building on interval-based uncertainty models, this dissertation develops a comprehensive interval-based integrity monitoring approach. It incorporates set-based fault detection and exclusion methods, along with the set-based central estimator. By adopting the multiple-hypotheses framework from ARAIM, the approach is capable of handling multiple simultaneous faults. This makes it particularly advantageous for urban navigation, where multiple unknown local threats may occur due to challenging environments.
- **Performance assessment of integrity and availability for aviation and urban navigation applications.** The developed methods are implemented and evaluated for aviation and urban navigation scenarios.
  - For **aviation users**, improvements to the baseline ARAIM algorithm are demonstrated, particularly in terms of reducing the computed fault-free Horizontal Protection Level (HPL) by incorporating error bounding for the least-squares estimator.
  - For **urban navigation**, the developed interval-based integrity monitoring approach is evaluated in two ways: (i) analytical and MC evaluation for integrity risk (IR) bounds, (ii) experimental evaluation and validation using real-world datasets collected from urban environments, with a focus on integrity and availability metrics.

Tab. 1.1 provides a summary linking the challenges to the corresponding contributions of this thesis. Readers of interest can directly refer to specific sections for further reading.

Thesis outline. The remainder of this thesis is structured in the following way:

Chapter 2 introduces the core concepts, models, and methods used in the thesis. It covers the fundamentals of interval and set, the concept of uncertainty, and a review of the navigation performance requirements. The chapter also provides a detailed overview of (Advanced) Receiver Autonomous Integrity Monitoring (RAIM) methods.

Chapter 3 is devoted to the distribution-free uncertainty modeling for GNSS pseudorange measurements. The residual tropospheric errors, residual ionospheric errors, and multipath errors are investigated, with corresponding methodologies and experimental validations discussed.

Chapter 4 focuses on state estimation and error bounding. This includes a revisit of error bounding for the classical least-squares estimator and a formal formulation of the set estimator, and the development of the set-based central estimator and the set-constrained least-squares

Table 1.1: Challenges of developing the interval-based framework for integrity monitoring and corresponding thesis contributions.

Challenges & Contributions	Section
Appropriate uncertainty modeling for GNSS observations:	
▸ Interval bounding for residual tropospheric errors	Sec. 3.2
▸ Interval bounding for residual tropospheric errors	Sec. 3.3
▸ Interval bounding for pseudorange multipath errors	Sec. 3.4
State estimation and corresponding error bounding:	
▸ Set estimator based on constraint satisfaction	Sec. 4.3.1
▸ Optimal set-based central estimator	Sec. 4.3.2
▸ Set-constrained least-squares estimator	Sec. 4.3.3
▸ Application: Improved fault-free Horizontal Protection Level (HPL) for aviation users using error bounding for the least-squares estimator	Sec. 4.2& Sec. 6.1
Fault detection and exclusion methods:	
▸ Set-based fault detection and its evaluation	Sec. 5.1
Integrity monitoring algorithms incorporating established uncertainty models:	
▸ Interval-based integrity monitoring architecture against multiple faults	Sec. 5.2
▸ Analytical evaluation for integrity risk (IR) bound through a benchmark example	Sec. 5.3.1& Sec. 5.3.2
▸ Monte Carlo evaluation for integrity risk (IR) bound under realistic scenarios	Sec. 5.3.3 Sec. 6.2
▸ Application: Real-world implementation and evaluation for urban navigation	

estimator, meeting different requirements. A comparative analysis of these methods is also provided.

Chapter 5 presents the set-based fault detection and exclusion method, based on which the interval-based integrity monitoring approach is developed. The proposed method is then compared to classical RAIM methods through analytical and MC evaluations.

Chapter 6 implements and evaluates the methods developed in this thesis, mainly in two experimental scenarios. The first one focuses on reducing HPL in the baseline ARAIM algorithm for aviation users based on error bounding for the least-squares estimator, and the second one applies the developed interval-based integrity monitoring approach in urban navigation scenarios.

The thesis concludes with a summary of the key findings and results in Chapter 7. Future research directions are also discussed, focusing on open questions and potential areas for further exploration.

2
Fundamentals

2.1 Basics of intervals and sets

Intervals and *sets* are fundamental mathematical concepts used throughout this thesis. This section introduces the definitions, notations, properties, and relevant concepts of interval mathematics and set theory.

2.1.1 Set operations and properties

Some fundamental concepts related to set operations and properties are essential in set theory. The following paragraphs provide an overview of these concepts, specifically in the context of general sets within Euclidean space.

Set operations

Considering two sets $\mathcal{X}_1, \mathcal{X}_2 \subseteq \mathbb{R}^m$, the main operations on sets are defined as follows:

- **Intersection**:
$$\mathcal{X}_1 \cap \mathcal{X}_2 \coloneqq \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{x} \in \mathcal{X}_1 \text{ and } \mathbf{x} \in \mathcal{X}_2\}. \tag{2.1}$$
- **Union**:
$$\mathcal{X}_1 \cup \mathcal{X}_2 \coloneqq \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{x} \in \mathcal{X}_1 \text{ or } \mathbf{x} \in \mathcal{X}_2\}. \tag{2.2}$$
- **Cartesian product**:
$$\mathcal{X}_1 \times \mathcal{X}_2 \coloneqq \{(\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{R}^{2\,m} \mid \mathbf{x}_1 \in \mathcal{X}_1 \text{ and } \mathbf{x}_2 \in \mathcal{X}_2\}. \tag{2.3}$$
- The **inclusion** of $\mathcal{X}_1$ in $\mathcal{X}_2$:
$$\mathcal{X}_1 \subseteq \mathcal{X}_2 \Leftrightarrow \forall \mathbf{x} \in \mathcal{X}_1,\ \mathbf{x} \in \mathcal{X}_2. \tag{2.4}$$
  In this equation, $\mathcal{X}_1$ is a subset of $\mathcal{X}_2$.
- **Equality**:
$$\mathcal{X}_1 = \mathcal{X}_2 \Leftrightarrow \mathcal{X}_1 \subseteq \mathcal{X}_2 \text{ and } \mathcal{X}_2 \subseteq \mathcal{X}_1. \tag{2.5}$$
- **Minkowski sum**:
$$\mathcal{X}_1 \oplus \mathcal{X}_2 := \{\mathbf{x}_1 + \mathbf{x}_2 \in \mathbb{R}^m \mid \mathbf{x}_1 \in \mathcal{X}_1, \mathbf{x}_2 \in \mathcal{X}_2\}. \tag{2.6}$$

Diameter and set boundedness

The *diameter* of a set $\mathcal{X} \subseteq \mathbb{R}^m$ is given by

$$\operatorname{diam}(\mathcal{X}) = \max_{\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}} \|\mathbf{x}_1 - \mathbf{x}_2\|. \tag{2.7}$$

A set is considered *bounded* (i.e., *a bounded set*) in a metric space if its diameter is finite (Hazewinkel, 1995).

Chebyshev center and radius

For a closed and bounded set $\mathcal{X} \subseteq \mathbb{R}^m$, the *Chebyshev center* of $\mathcal{X}$ is defined as

$$\operatorname{cen}(\mathcal{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^m} \max_{\mathbf{x}_\mathcal{X} \in \mathcal{X}} \|\mathbf{x} - \mathbf{x}_\mathcal{X}\|. \tag{2.8}$$

The *Chebyshev radius* of $\mathcal{X}$ is defined as

$$\operatorname{rad}(\mathcal{X}) = \max_{\mathbf{x}_\mathcal{X} \in \mathcal{X}} \|\operatorname{cen}(\mathcal{X}) - \mathbf{x}_\mathcal{X}\|. \tag{2.9}$$

From the previous definitions, it follows that (Casini, 2002)

$$\operatorname{rad}(\mathcal{X}) \leq \operatorname{diam}(\mathcal{X}) \leq 2 \operatorname{rad}(\mathcal{X}), \text{ with } \mathcal{X} \subseteq \mathbb{R}^m. \tag{2.10}$$

Geometrically, $\operatorname{cen}(\mathcal{X})$ is the center of the smallest circumscribing (hyper)sphere of $\mathcal{X}$, and its radius is $\operatorname{rad}(\mathcal{X})$.

Literature such as Amir (1978) and Casini (2002) adopt various norm operations for the above definitions. In the context of this thesis, the norm $\|\cdot\|$ is restricted to ℓ_2 (Euclidean distance), sufficient for GNSS positioning and navigation problems. Subsequently, the investigated Chebyshev center yields the following properties:

- **Existence**: The Chebyshev center always exists;
- **Uniqueness**: The Chebyshev center is unique.

2.1.2 Polytopes

Two definitions of polytopes are reported here, which are mathematically equivalent (Ziegler, 1995):

*An H-**polytope** denotes a bounded intersection of closed halfspaces:*

$$\mathcal{P}(\mathbf{H}, \mathbf{b}) = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{H}\,\mathbf{x} \leq \mathbf{b}, \mathbf{H} \in \mathbb{R}^{n \times m}, \mathbf{b} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m. \tag{2.11}$$

Eq. 2.11 can be interpreted as the solution set of a finite number of linear inequalities.

*A V-**Polytope** denotes a convex hull of a finite point set:*

$$\mathcal{P} = \operatorname{conv}(\mathbf{V}) \subseteq \mathbb{R}^m, \quad for \quad \mathbf{V} \in \mathbb{R}^{m \times u}. \tag{2.12}$$

In Eq. 2.12, u points represent its vertices, and the operation $\operatorname{conv}(\cdot)$ forms the convex hull.

Some useful properties of polytopes include:

- Every intersection of a polytope with an affine subspace is a polytope;
- The Minkowski sum of two polytopes is a polytope;
- Every projection of a polytope is a polytope.

There are several ways to define the center of a polytope, for example, the following concepts:

- The **center of gravity**, or **centroid**;
- Mean position of all vertices, termed as **vertex centroid**, or **barycenter**;
- Point at the location where the product of distances from all boundaries (hyperplane) is maximized, i.e., **analytic center**;
- The center of the biggest (hyper-)sphere inside the polytope, i.e., **incenter**;
- The center of the smallest (hyper-)sphere containing the polytope, i.e., the **Chebyshev center**.

For more details, refer to textbooks such as Ziegler (1995) and Boyd and Vandenberghe (2004), or the article by Barnes and Moretti (2005), which provides a comprehensive summary of polytope centers.

2.1.3 Zonotopes

Zonotopes form a special class of polytopes. Each zonotope is a convex polytope that is symmetric with respect to its center. They are utilized in many technical applications and from different viewpoints. Ziegler (1995) gives the definition of zonotopes:

> *A zonotope $\mathcal{Z}(\mathbf{c}_\mathcal{Z}, \mathbf{G}_\mathcal{Z})$ is defined as the image of a hypercube under an affine projection:*

$$\mathcal{Z}(\mathbf{c}_\mathcal{Z}, \mathbf{G}_\mathcal{Z}) = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{x} = \mathbf{c}_\mathcal{Z} + \sum_i^n w_i \cdot \mathbf{g}_\mathcal{Z}^{\{i\}}, -1 \leq w_i \leq 1\}.$$

In Eq. 2.1.3, $\mathbf{c}_\mathcal{Z}$ is the center of the zonotope, $\mathbf{g}_\mathcal{Z}^{\{i\}}$ are $m-$dimensional column vectors referred to as generators, and $\mathbf{G}_\mathcal{Z}$ is the $m \times n$ generator matrix. The generators of a zonotope determine its shape relative to its center.

Alternatively, the H-zonotope can be described as the intersection of n pairs of parallel hyperplanes (Eppstein, 1995):

$$\mathcal{Z}(\mathbf{H}, \mathbf{b}) = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{H}\,(\mathbf{x} - \mathbf{c}_\mathcal{Z}) \leq \mathbf{b}\}. \tag{2.13}$$

The shape of zonotopes is invariant under rotations, meaning that the orientation of the coordinate system does not affect the shape of the zonotope or associated scalar measures. In addition, zonotopes are suitable measures of point uncertainty for linear estimators (Schön and Kutterer, 2005b). This feature will be further elaborated from the perspective of interval mathematics in the subsequent paragraphs. For more information on zonotopes, readers of interest can refer to Ziegler (1995) and Henk et al. (1997).

2.1.4 Intervals

The basics of interval notations, computations, and functions are given below. For further reading on interval mathematics and its applications, readers of interest are referred to textbooks such as Jaulin et al. (2001) and Moore et al. (2009).

Definitions and notations

In this work, an *interval* is denoted by

$$[x] \coloneqq [\underline{x}, \overline{x}] = \{x \in \mathbb{R} \mid \underline{x} \leq x \leq \overline{x}\}, \tag{2.14}$$

which is a *closed* and *connected* subset of $\mathbb{R}$. $\underline{x}$ is the lower bound (left endpoint) and $\overline{x}$ is the upper bound (right endpoint). An interval is said to be bounded if both its lower and upper bounds are finite real values. If $\underline{x} = -\infty$ or $\overline{x} = \infty$, $[x]$ is an unbounded interval.

There is some debate on whether the empty set $\emptyset$ should be considered an interval (Jaulin et al., 2001). Despite the ambiguity in defining its bounds, this thesis considers the empty set as an interval. This ensures that the set of intervals remains closed with respect to intersections, and this is beneficial in applications where $\emptyset$ represents the absence of a solution.

The basic properties of a non-empty and bounded interval are as follows:

- **Midpoint/center:**
$$\mathrm{mid}([x]) = x_m \coloneqq \frac{\underline{x} + \overline{x}}{2};$$
- **Width:**
$$\mathrm{wid}([x]) = x_w \coloneqq \overline{x} - \underline{x};$$
- **Radius:**
$$\mathrm{rad}([x]) = x_r \coloneqq \frac{x_w}{2} = \frac{\overline{x} - \underline{x}}{2}.$$

Special cases to consider:

- **Degenerate Intervals ($\underline{x} = \overline{x}$):** In this scenario, an interval is considered to *degenerate* into a real value, which makes properties like width and radius irrelevant. Inversely, any real number can be viewed as a degenerate interval in interval computations;
- **Unbounded Intervals ($\underline{x} = -\infty$ and/or $\overline{x} = \infty$):** In this case, an interval is considered unbounded and is not closed, with $\mathrm{wid}([x]) = \infty$. This case is not considered in this thesis due to its limited physical significance in uncertainty evaluation.

An n-dimensional interval vector is an ordered n-tuple of intervals:

$$[\mathbf{x}] \coloneqq \begin{bmatrix} [\underline{x}_1, \overline{x}_1] \\ \cdots \\ [\underline{x}_n, \overline{x}_n] \end{bmatrix}, \ \mathbf{x} \in \mathbb{R}^n. \tag{2.15}$$

$[\mathbf{x}]$ can be interpreted as an Cartesian product of n intervals. In this regard, $[\mathbf{x}]$ is also named an interval box, which is a subset of $\mathbb{R}^n$. Each component $[x_i]$ is the projection of $[\mathbf{x}]$ onto the i-th axis. Specially, the projection of an empty set of $\mathbb{R}^n$ onto any axis is an empty interval $\emptyset$. Fig. 2.1 illustrates a two-dimensional interval box with its projections onto the axes resulting in intervals $[\underline{x}_1, \overline{x}_1]$ and $[\underline{x}_2, \overline{x}_2]$.

For simplicity, this thesis adopts the notation for an interval vector: $[\mathbf{x}] = [\underline{\mathbf{x}}, \overline{\mathbf{x}}]$ with $\underline{\mathbf{x}}, \overline{\mathbf{x}} \in \mathbb{R}^n$, i.e., using vectors of upper bounds and lower bounds.

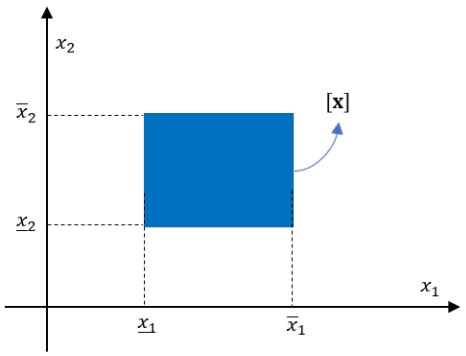


Figure 2.1: An example two-dimensional interval box with its interval components $[\mathbf{x}] = [x_1] \times [x_2]$.

Interval arithmetics

The real-valued arithmetic operations can be mostly extended to intervals, with the same basic operators: $+$, $-$, $\times$, $\div$:

$$[x] \diamond [y] = [x \diamond y \in \mathbb{R} \mid x \in [x], y \in [y]], \tag{2.16}$$

where $[x]$ and $[y]$ are intervals and $\diamond$ can be any of the operations listed above.

In addition, as intervals can essentially be viewed as one-dimensional sets, the standard set operations also apply to interval values, cf. Sec. 2.1.1.

Interval functions

Elementary functions such as sin, tan, *exp*, etc., extend to intervals. If f represents a real-valued function of a single real variable x (i.e., from $\mathbb{R}$ to $\mathbb{R}$), the range of $f(x)$ as x varies within an interval $[x]$ is defined as the image of the set $\{x \mid x \in [x]\}$ under the mapping f:

$$f([x]) = \{f(x) \mid x \in [x]\}. \tag{2.17}$$

Analogously, for a function f of multiple real variables $\mathbf{x} = [x_1, ..., x_m]^T$, the output interval when the i-th variable x_i varies within an interval $[x_i]$ can be expressed as

$$f([x_i] \mid \mathbf{x}^*) = \{f(\mathbf{x}) \mid x_i \in [x_i], \mathbf{x}_{(...,i-1,i+1,...)} = \mathbf{x}^*_{(...,i-1,i+1,...)}\}. \tag{2.18}$$

Eq. 2.18 represents the set image of $[x_i]$ under the mapping f, with all other variables taking values from $\mathbf{x}^*$. This reflects the actual range of the function's output as the variable x_i changes and can be computed using interval arithmetic. For illustration, a numerical example is given in Fig. 2.2 for a bivariate function $f(x_1, x_2) = x_1^2 + x_2$.

Consider a general case of a function $\mathbf{f}$ from $\mathbb{R}^n$ to $\mathbb{R}^m$. An interval counterpart, denoted by $[\mathbf{f}]([\mathbf{x}])$, is an *inclusion function* for $\mathbf{f}$ if

$$\forall [\mathbf{x}] \in \mathbb{IR}^n, \ \mathbf{f}([\mathbf{x}]) \subseteq [\mathbf{f}]([\mathbf{x}]), \tag{2.19}$$

where $\mathbb{IR}^n$ denotes the set of all n-dimensional boxes. An example of $\mathbf{f}$ from $\mathbb{R}^3$ to $\mathbb{R}^2$ is given in Fig. 2.3. The set image of $[\mathbf{x}]$ under $\mathbf{f}$ is displayed as the blue hexagon in Fig. 2.3(b), representing the factual range $\mathbf{f}([\mathbf{x}])$; as a comparison, the gray box represents the minimal inclusion function $[\mathbf{f}]([x])$, overestimating the variation of $\mathbf{f}([\mathbf{x}])$. This phenomenon is conventionally termed *wrapping effect*.

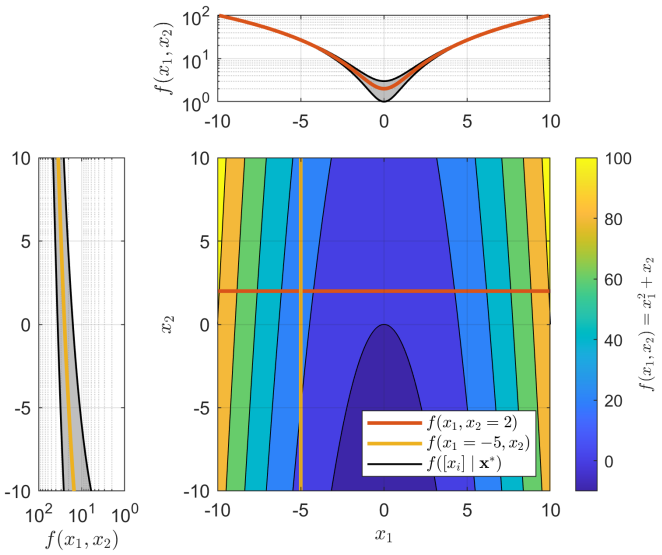
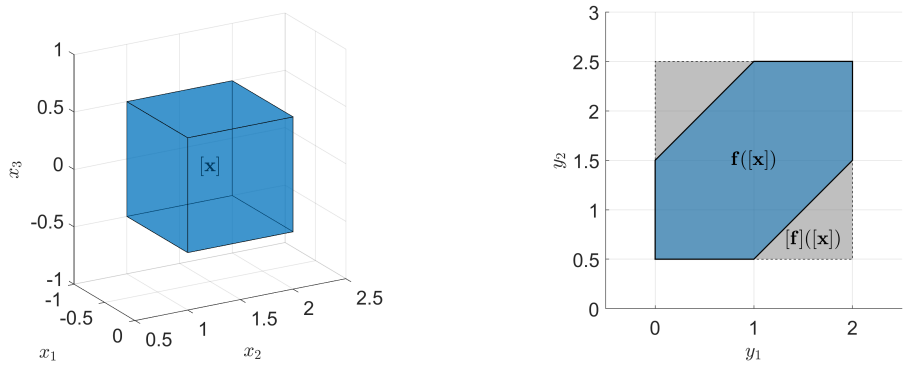


Figure 2.2: An illustrative numerical example for function $f(x_1, x_2) = x_1^2 + x_2$. Function values varying with x_1 and x_2 are shown as contour plots in the lower right image. Cross-sections of $x_1 = -5$ (yellow) and $x_2 = 2$ (red) are chosen and displayed in the left and upper images separately. The range of $f(x_1, x_2)$ as x_1 varies within $[-6, -4]$, and the one as x_2 varies within $[1, 3]$, can be computed by corresponding set images under the mapping f. The results are shown as gray areas, respectively.



(a) $\mathbf{x} \in \mathbb{R}^3$ (b) $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in \mathbb{R}^2$

Figure 2.3: An illustrative numerical example for function $\mathbf{f}: \mathbf{y} = \mathbf{A}\,\mathbf{x}$ from $\mathbb{R}^3$ to $\mathbb{R}^2$. The input interval $[\mathbf{x}] = \begin{bmatrix} [-1, 0] \\ [1, 2] \\ [-0.5, 0.5] \end{bmatrix}$ (shown as the blue cube in (a)), matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, the set image $\mathbf{f}([x])$ and inclusion function $[\mathbf{f}]([x])$ (shown as the blue and gray areas in (b)).

The same principle applies to constructing a zonotope. The definition presented in Eq. 2.1.3 can be understood as the set image of an m-dimensional interval box $[\mathbf{x}] \in \mathbb{R}^m$ under an affine projection $\mathbf{f}$. In this context, the actual range $\mathbf{f}([\mathbf{x}])$ forms a zonotope, while its interval bounding box belongs to the family of $[\mathbf{f}](\mathbf{x})$. The example illustrated in Fig. 2.3 represents a specific case of a zonotope with three generators. In this regard, using a zonotope as a measure of uncertainty avoids the possibility of overestimation.

2.2 Concept of uncertainty

Qualifying measurement results is a critical aspect of any experimental approach (Ferrero and Salicone, 2006). The International Vocabulary of Metrology (VIM) and Guide to the

expression of uncertainty in measurement (GUM), representing the most important reference documents in metrology, accept the concept that the uncertainty of measurement characterizes the dispersion of the measured values that could be reasonably attributed (JCGM 100:2008, 2008; JCGM 200:2012(E/F), 2012). GUM interprets uncertainty as a measure of the potential error in the measurement result, clearly distinguishing it from the term "error", which is explicitly defined as the difference between true and measured values. Errors, however, can be unknown and unknowable and can hardly be determined with confidence due to the challenges of obtaining a sufficiently "true" value. For example, the measurement result can unknowingly be very close to the true value, resulting in a negligible error, yet may carry a large uncertainty. Additionally, the uncertainties associated with the random and systematic effects that give rise to the error can be evaluated. Therefore, the concept of uncertainty is often more meaningful and valuable in various situations. This thesis adopts the aforementioned GUM traditions, emphasizing the clear distinction between 1) the measurement result and its evaluated uncertainty and 2) the truth and error.

2.2.1 Uncertainty component classification

In GNSS applications, random variability (i.e., stochasticity) and imprecision are two prominent sources of measurement errors (Kutterer, 2002):

- *Random error* arises from unpredictable or stochastic temporal and spatial variations of influence quantities, termed *random variability*. The uncertainty due to random variability may be characterized by a standard deviation. Such a standard deviation can be evaluated either from experimentally observed frequency distributions or from assumed probability distributions based on experience or other comparatively reliable information. These are categorized as *Type A* and *Type B* evaluations, respectively, according to GUM.
- A recognized systematic effect may be quantified and compensated by a correction model so that the subsequent estimator involved is not biased. However, due to the inherently unknowable nature of errors, the deterministic correction may only approximate the true value, reflecting our current state of knowledge. Consequently, *remaining systematic errors* may persist due to the imperfect correction of systematic effects. Any assumptions deviating from reality and approximations made in the correction model can lead to uncertain corrected values. Additionally, input quantities to the correction model, such as constants and parameters, are often only imprecisely known or not fully representative over the observation period, which can introduce additional uncertainty. All these factors contribute to *imprecision* in the measurements.

At this point, it has been argued that errors arise from both random and systematic effects, and these errors are recognizable and associated with various uncertainty components. However, categorizing these components strictly as "random" or "systematic" remains inconclusive. GUM recommends that the distinction between random and systematic uncertainty components should not exclusively depend on whether they originate from random or systematic effects but rather on the evaluation method applied, as outlined in Recommendation INC-1 by the Working Group on the Statement of Uncertainties (Kaarls, 1981). This approach minimizes ambiguity in categorizing uncertainty components and emphasizes the importance of assessing the overall uncertainty budget (Kutterer, 2001). For instance, the result of one measurement might be used as an input datum to the second measurement, whereas a "systematic" component of uncertainty in the latter could stem from a "random" component in the former. Consequently, the method of combing uncertainties must be carefully designed to meet the safety requirement, which is to be addressed in the following sections.

2.2.2 Uncertainty representations

It is common practice to assume that GNSS measurements have been corrected for all recognized significant effects, and the remaining errors are considered as random variability, e.g., treated as Gaussian distributed. This makes variance or standard deviation a meaningful representation of uncertainty, which is convenient from a methodological perspective. Currently, this approach is predominantly adopted in GNSS applications.

However, a purely stochastic handling of the GNSS uncertainty is not always adequate in the following scenarios:

1. *For complex systems where dominant systematic error components are observed, treating them as purely random variables may lead to unsatisfactory results* (Milanese and Vicino, 1991; Schön et al., 2018).

   This is not uncommon because of incomplete knowledge about quantifying various effects, limited hardware for onboard computations, economic considerations, or particular user requirements. For instance, nominal signal deformations caused by differences in the signal-generating hardware onboard the GNSS satellites can cause range errors that vary with user receivers and are reported non-Gaussian behaved. They are modeled as components of nominal biases, upper bounded by constant values that are input to the onboard integrity monitoring algorithms (Macabiau et al., 2015). Sec. 2.4.2 gives a detailed introduction.

2. *The considered error component may take any value between bounds, but there is an insufficient basis for assuming distributions due to incomplete prior knowledge* (Beer et al., 2014).

   For example, pseudorange multipath errors cannot be effectively handled by stochastic methods, as their magnitude is influenced by various factors and cannot be reduced by averaging repeated observations due to their non-zero mean. Nevertheless, multipath error envelope can be used to quantify their upper and lower bounds, within which the actual error magnitude oscillates due to changes in the relative phase of the multipath. This topic is discussed in further detail in Section 3.4.

   Another example of uncertainty arises from the limited resolution of a digital instrument, as noted by GUM. If we denote the resolution as δ, a range of inputs to the measurement instrument may yield the same indication y of finite digits. Even when identical measurements are repeated, the uncertainty due to repeatability is not zero. The expected distribution of the measurement results falls within the range $[y-\delta/2, y+\delta/2]$. However, this distribution is unknown and cannot be observed through repeated measurements. Similarly, rounding or truncating numbers during automated data processing can introduce uncertainty in real-valued data, which can be difficult to predict (JCGM 100:2008, 2008). In this thesis, both scenarios are referred to as *rounding errors*, addressed in the subsequent subsection.

3. *Parameters that are set-valued, with any real-valued candidate being acceptable, require interval representation.*

   For instance, the Power Spectral Density (PSD) bounding approach can handle the residual tropospheric error (Gallon et al., 2021b) and Inertial Measurement Unit (IMU) errors (Gallon et al., 2021a) by establishing upper bounds for the historical observation data in the PSD domain. Certain parameters may be set-valued due to mathematical interdependencies within the nonlinear model. Uncertainty can arise from the multiple choices of parameter values in the applications (Su et al., 2024).

4. *Assumptions about the distribution of the considered error component may be difficult to validate, or even unrealistic, but are critical to the system's reliability* (Combastel, 2003).

   In aviation integrity studies, Gaussian distribution assumptions for GNSS pseudorange measurements are untrue at very low probabilities due to limitations of the Central Limit Theorem (CLT), i.e., the observed probability distributions often have tails fatter than a standard normal distribution. Overbounding methods have been developed to address these issues, retaining the usage of Gaussian, nonetheless compromising over-pessimistic confidence levels (DeCleene, 2000; Schempp and Rubin, 2002; Rife et al., 2004a; Blanch et al., 2019b).

These concerns particularly motivate the investigation and application of interval representation to denote uncertainty measures. Intervals can capture stochasticity by indicating confidence levels or represent worst-case scenarios for imprecision. Fig. 2.4 displays the conceptual model of an uncertainty interval. The red arrow indicates the true value of the quantity of interest, while the probability distribution along the upper axis quantifies the frequency of specific values occurring within the interval. Another interpretation, shown in the lower axis, refers to an interval range encompassing all possible measurement results, regardless of the distribution or without any distributional assumptions. Ultimately, the uncertainty of a measurement is conveyed through an interval placed around the measured value, with the belief that the true value lies within this range. Therefore, this approach does not necessarily involve any stochastic information.

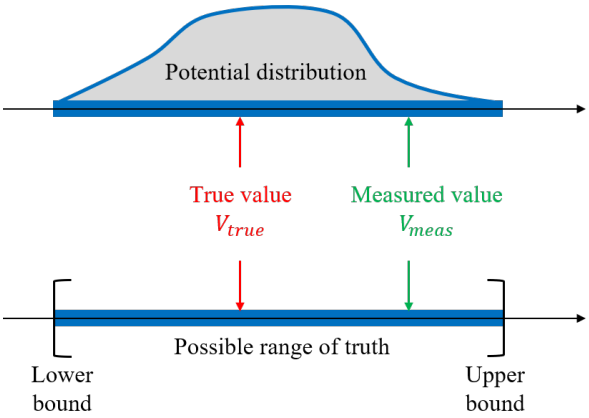


Figure 2.4: Conceptual model of uncertainty intervals.

This approach is natural, straightforward, and has been applied across various engineering disciplines for decades (Kutterer, 1994; Kreinovich, 1995; Jaulin et al., 2001; Althoff, 2010; Cimino et al., 2014; Rohou et al., 2020; Kabir et al., 2024). Rather than attempting to minimize the impact of a specific systematic effect, it offers a realistic description of uncertainty (Schön and Kutterer, 2006). The GUM also recognizes this perspective in its discussion of "expanded uncertainty", defined as:

> *Quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measured quantity* (Sec. 6.1.2 in JCGM 100:2008 (2008)).

The interval representation adopted in this work aligns with the scope of "expanded uncertainty", which is intended for particular commercial, industrial, and regulatory applications. In this light, it offers a practical and comprehensive framework for uncertainty evaluation in GNSS applications, where both random variability and remaining systematics coexist.

2.2.3 Variance-based evaluation principles

GUM introduces two ways of evaluating uncertainty components, i.e., *Type A* and *Type B* evaluations, in terms of stochastic representations:

- ***Type A.*** The variance or standard deviation is estimated from a series of repeated observations. It is derived from a probability density function based on an observed frequency distribution.
- ***Type B.*** The variance or standard deviation is evaluated using available knowledge. It is based on an assumed probability density function that reflects the degree of belief regarding the occurrence of an event.

At this point, all uncertainty components are treated stochastically, and the law of quadratic variance propagation is applied to the method of sensitivity analysis for the measurement model, expressed as follows:

$$y = f(\mathbf{d}) \tag{2.20}$$

In this equation, y represents the estimate of the measured quantity, whose uncertainty depends on the values of input quantities $d_1, d_2, ..., d_N$ that are fed into the measurement model f. The vector $\mathbf{d} = [d_1, d_2, ..., d_N]^T$ contains exhaustive, potentially countless influence factors for the corresponding measured quantity and measurement model.

Subsequently, the variance (denoted by σ_f^2), an uncertainty measure for the estimate of the measured quantity, can be determined using the following equation:

$$\sigma_f^2 = \sum_i \left[\frac{\partial f(\mathbf{d})}{\partial d_i}\right]^2 \sigma_{d_i}^2, \tag{2.21}$$

where

- $\mathbf{d}$ includes all influence factors that contribute to measurement uncertainty, which are assumed to be statistically independent;
- $\sigma_{d_i}^2$ represents the variance related to the uncertainty component associated with influence factor d_i. This variance is evaluated using either *Type A* or *Type B* approach;
- $\frac{\partial f(\mathbf{d})}{\partial d_i}$ refers to sensitivity coefficients that describe how the output estimate varies with changes in the input values.

The underlying concept is as follows. Typically, the quantity of interest is not measured directly but is instead derived from N other quantities through a functional relationship known as the measurement model. Consequently, the uncertainty associated with the measurement result (represented by the estimated standard deviation of the output estimate) depends on the estimated standard deviations of each input estimate (influence factors). Furthermore, each uncertainty component may be influenced by additional factors, which can introduce sub-level uncertainty components, thereby leading to a complex functional model.

In GNSS applications, the functional relationship can be modeled by the non-linear GNSS observation equation, cf. Sec. 3.1. The uncertainty of the pseudorange measurements after correcting all recognized systematic effects can be evaluated through Eq. 2.21.

Another interpretation can be found in the theory of *elementary error model (EEM)*. Each realization of a measured quantity differs from its expected value by a random deviation, presented as a sum of numerous very small elementary errors (Pelzer, 1985; Schwieger, 1999). The individual elementary errors are assumed to have the same absolute value, with equal probabilities for both negative and positive signs. Thus, the expected value of the random

deviation is zero. As the analysis of a measurement process gets more detailed, the number of elementary errors increases while their absolute values decrease. Assuming an infinite number of elementary errors with infinitely small absolute values justifies the use of a standard normal distribution for the standardized random deviation. The influence of the elementary errors on the measurement should be modeled by influence factors, which can be integrated into influence matrices and analyzed through sensitivity analysis. For instance, Schwieger (2007a) illustrates this approach using GPS data to detect tectonic movements in Romania, while Kerekes and Schwieger (2020) applies it to laser scanner data for the Kops arch dam in Austria.

In the case of correlated influence factors, Eq. 2.21 should be adapted to account for interdependence, which is beyond the scope of this work. Readers of interest are referred to GUM series (ISO/IEC GUIDE 98-3:2008(E), 2008; JCGM 100:2008, 2008; JCGM 102:2011, 2011; JCGM GUM-6:2020, 2020) and literature on EEM (Pelzer, 1985; Schwieger, 1999, 2007a; Kerekes and Schwieger, 2020) for further information.

2.2.4 Interval-based evaluation principles

Kutterer (1999) introduces a method to determine interval-presented uncertainty using sensitivity analysis based on interval mathematics. This method has been applied to geodetic monitoring network (Schön and Kutterer, 2001b; Neumann et al., 2006) and to GPS data analysis (Kutterer, 2001; Schön and Kutterer, 2005a; Dbouk, 2021).

The fundamental concept is that the empirical models are built on assumptions and approximations with inherent uncertainty. Constants and parameters in these models are often only imprecisely known and are associated with a range of values. These uncertainties contribute to a maximum interval range of variation for the model's output, which reflects the uncertainty due to remaining systematics. This approach differs from the GUM sensitivity analysis, as it discards stochastic assumptions about systematic errors.

Consider a functional model f with N influence factors $\mathbf{d} = \{d_1, d_2, ..., d_N\}$. Given nominal values $\mathbf{d}^*$ and associated uncertainties represented by interval radii $\mathbf{d}_r$, the matrix of sensitivity coefficients $\mathbf{F}$ and resulting interval radius f_r of the output estimate can be determined as follows:

$$\mathbf{F} = \frac{\partial f(\mathbf{d}^*)}{\partial \mathbf{d}}, \;\; f_r = |\mathbf{F}| \cdot \mathbf{d}_r, \tag{2.22}$$

assuming that the influence factors are independent. In Eq. 2.22, $\mathbf{F}$ denotes the matrix of partial derivatives evaluated at $\mathbf{d}^*$, and $|\mathbf{F}|$ refers to the element-wise absolute values.

Unlike Eq. 2.21, which is based on variance-based representations, Eq. 2.22 propagates uncertainty through partial differentiation in an interval framework. The uncertainties in the influence factors are modeled as intervals centered at $\mathbf{d}^*$ with radii $\mathbf{d}_r$, resulting in an interval output estimate bounded by $[-f_r, f_r]$.

Undoubtedly, the key to the interval-based sensitivity analysis is the reliable characterization of the model's influence factors and the assessment of their uncertainties. Schön and Kutterer (2006) categorizes influence factors for practical applications into the following categories:

- **Model constants** are often determined empirically and should be associated with an uncertainty interval that reflects their quality;
- **Model parameters** are input values to adapt the model to specific observational conditions. They typically represent average values along with associated uncertainty measures;

- **Auxiliary information** includes any external source data, products, or information necessary for the model that can influence the uncertainty of the model's output. The uncertainty intervals for this information must be evaluated based on the quality indicators and incorporated into the sensitivity analysis.

Consistently, *Type A* and *Type B* evaluations are effective for supporting interval-based sensitivity analysis:

- ***Type A.*** The interval is assessed based on repeated observations, which could include the maximum range or specific quantiles of the observed values.
- ***Type B.*** The interval is determined using available knowledge, specifically focusing on possible error bounds.

For *Type B* evaluation, GUM depicts a pool of available information. This information aids in evaluating model influence factors through scientific judgment. Key sources of uncertainty information include:

- Uncertainty information indicated during the model construction process;
- Expert knowledge or experience regarding the behavior and properties of relevant factors;
- Manufacturer's specification;
- Information provided by model's accuracy evaluations;
- Uncertainties associated with reference data taken from established handbooks.

As an example, *rounding errors*, as introduced earlier, arise from two main sources: (i) the limitations in the resolution of the instrument's display and (ii) the rounding or truncation of numbers during data processing. The associated interval-described uncertainty is readily determined by the instrument's specified resolution and the number of decimal places reported. For example, if a manufacturer specifies a resolution of 10^{-3} or if digital readings are to three decimal places, then an interval with a radius of 0.5×10^{-3} captures the uncertainty due to limited resolution. This serves as the minimum standard for assessing the uncertainty of any quantities in this work, particularly in the absence of additional information.

In Chap. 3, the interval-based sensitivity analysis method is revised and extended to fulfill practical requirements.

2.2.5 Terminology in other disciplines

Terminology related to errors and uncertainty differs across various scientific disciplines, resulting in different classifications and conceptualizations. The emphasis here is not on assessing correctness or accuracy but rather on highlighting the scientific and philosophical approaches that are appropriate for different contexts. Below are some examples with brief explanations:

- **Statistical and systematic uncertainties**: These are commonly encountered in physics-related fields like cosmology (Conrad et al., 2003) and nuclear/particle physics (Pérez et al., 2016).
- **Aleatory (or aleatoric) and epistemic uncertainties**: These terms are frequently used across several disciplines, particularly in the engineering reliability community (Oberkampf et al., 2004; Möller and Beer, 2008; Faes et al., 2021) and deep learning/-machine learning community (Abdar et al., 2021; Hüllermeier and Waegeman, 2021).

- **Objective or subjective uncertainties**: These uncertainties may be linked to human factors or environmental influences in ethics (Tannert et al., 2007), decision-making (Campos et al., 2007), psychology (Alquist and Baumeister, 2023), etc.
- **Reducible or irreducible uncertainties**: This classification appears in various papers and can complement other terminology. For example, Campos et al. (2007) differentiates objective from subjective uncertainties based on whether they can be principally reduced through additional empirical efforts. Similarly, Hüllermeier and Waegeman (2021) interprets aleatory uncertainties as irreducible and epistemic uncertainties as reducible within the total uncertainty budget.
- **Substantive and procedural uncertainties**: These terms are used in studies on economic behavior to describe the origins of uncertainties arising from human actions (Dosi and Egidi, 1991).
- **Data and knowledge uncertainties**: These concepts are typically introduced in deep learning and artificial intelligence to explicitly indicate the sources of uncertainties that arise when there is a mismatch between training and testing data (Abdar et al., 2021).

Uncertainty classification can often be ambiguous even within the same field, as authors may use interchangeably referenced terminologies and different concepts simultaneously. For instance, data/knowledge uncertainties and aleatory/epistemic uncertainties are frequently accepted in machine learning studies. Therefore, maintaining clarity in conceptualization is crucial and should be consistently emphasized.

2.3 Navigation performance requirements

The goal for most satellite navigation users is *assured* positioning, navigation and timing (PNT). Combining dissimilar sources of basic PNT information can lead to a much more robust positioning capability, i.e., greater PNT assurance (Parkinson et al., 2021). The United States Federal Aviation Administration (FAA) specifies four criteria to measure navigation performance: *accuracy, integrity, continuity*, and *availability.*

Accuracy. Accuracy is the simplest and most commonly used performance metric of a navigation system. It is a measure of the position error, i.e., the difference between the estimated position and the true position, that will be experienced by a user with a certain probability at any given instant. Particularly, the probability used for accuracy requirements in civil aviation is 95% probability, as specified by ICAO, for different operations. For example, the en-route horizontal accuracy requirement is 3.7 km, meaning that the horizontal position error is no greater than 3.7 km at 95% probability under nominal conditions.

Integrity. In addition to providing a PNT function,

> *a GNSS system must be able to provide valid and timely warnings to the user when the system must not be used for the intended operation.*

This capability is known as the integrity of the system. Integrity is a measure of the trust that can be placed in the correctness of the information supplied by the total system.

The metric typically used to quantify integrity is *integrity risk (IR)*, which concerns both the position error bounds and the required time to alert the user if that position bound cannot be assured to the required level of confidence. Algorithms in the user equipment compute a

Protection Level (PL) each time a position is estimated. If the computed PL exceeds a required *Alert Limit (AL)* for the operation, an alert is raised within the *Time-to-Alert (TTA)*. The IR requirement used in aviation is generally set to a probability of 10^{-7}, meaning that the probability of a Position Error (PE) exceeding the AL without warning to the user is 10^{-7}, or, alternatively, the probability of the position error is smaller than the AL is 99.99999%.

Integrity is essential to safety-critical applications such as civil aviation, and thus, it is a primary driver behind the development of various augmentation systems. Ground-based Augmentation System (GBAS) and Satellite-based Augmentation System (SBAS) are standardized GNSS augmentation systems that have been deployed to support civil aircraft navigation. In addition, an Aircraft-based Augmentation System (ABAS) augments and/or integrates information from the other GNSS elements with information available onboard the aircraft. RAIM is a basic type of ABAS that takes advantage of redundant information from the GNSS constellation itself to verify integrity. This will be further introduced in Sec. 2.4.

The Stanford diagram is a popular visualization tool for integrity performance assessment, as a two-dimensional histogram presenting the relationship of PE, PL, and AL for a set of measurements, cf. Fig. 2.5. It partitions the entire space into four categories:

- **Normal Operation (NO):** PE<PL<AL;
- **Misleading Information (MI):** PL<PE<AL, and AL<PL<PE;
- **Hazardous Misleading Information (HMI):** PL<AL<PE;
- **System Unavailable (SU):** AL<PL.

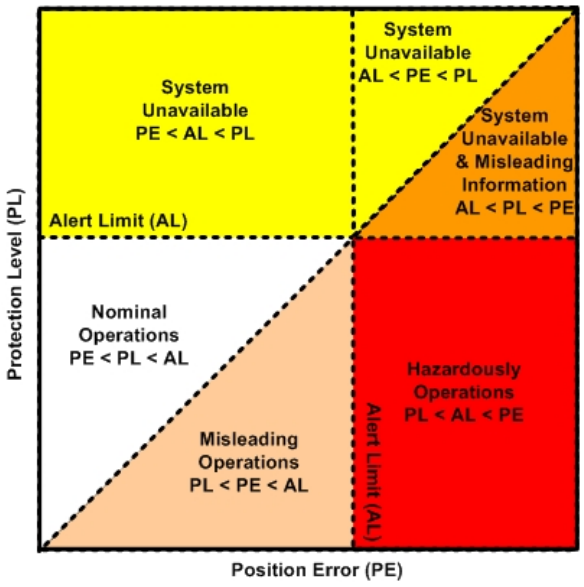


Figure 2.5: An example of the Stanford diagram.

Notably, the necessity of ground truth for computing PE prevents this approach in real-world applications. Nevertheless, it is applicable to simulations or experiments with accessible, high-quality reference solutions and, hence, is an effective way for evaluating integrity monitoring algorithms.

Continuity. International Civil Aviation Organization (ICAO) defines the continuity of a service as

> *the capability of the system to perform its function without unscheduled interruptions during the intended operation.* (ICAO, 2023)

Specifically, the navigation system must meet the accuracy and integrity requirements for the duration of the operation at the required probabilities. This is usually quantified by the *continuity risk (CR)*, which is defined as the probability of a detected but unscheduled navigation function interruption after the operation has been initiated. For instance, the continuity requirement is specified over a time period of 15s for approach operations that are relatively short.

Availability. Availability is defined as

> *the probability that a user is able to determine its position with the specified accuracy and is able to monitor the integrity of its determined position at the initiation of the intended operation.*(ICAO, 2023)

ICAO gives a range of values for the availability requirements concerning specific operations, which depend on several factors and should be selected accordingly.

Tab. 2.1 showcases the current navigation performance requirements for different aircraft operations by ICAO.

Table 2.1: Navigation performance requirements for civil aviation operations (ICAO, 2023).

Operation	**Accuracy** (horizontal, vertical)	**Integrity Risk (IR)**	**Time-to-Alert**	**Continuity Risk (CR)**	**Availability**
En-route	3.7 km, N/A	10^{-7}/h	5 min	10^{-4}/h to 10^{-8}/h	0.99 to 0.99999
En-route, Terminal	0.74 km, N/A	10^{-7}/h	15 s	10^{-4}/h to 10^{-8}/h	0.99 to 0.99999
Non-precision approach (NPA)	220 m, N/A	10^{-7}/h	10 s	10^{-4}/h to 10^{-8}/h	0.99 to 0.99999
APV-I*	16 m, 20 m	2×10^{-7}/h in any approach	6 s	8×10^{-6}/h per 15 s	0.99 to 0.99999
APV-II*	16 m, 8 m	2×10^{-7}/h in any approach	6 s	8×10^{-6}/h per 15 s	0.99 to 0.99999
Category I precision approach	16 m, 6 m to 4 m	2×10^{-7}/h in any approach	6 s	8×10^{-6}/h per 15 s	0.99 to 0.99999

Operation	**Alert Limit (AL)**	
	Horizontal	Vertical
En-route		
(oceanic/continental low density)	7.4 km	N/A
(continental)	3.7 km	N/A
En-route, Terminal	1.85 km	N/A
Non-precision approach (NPA)	556 m	N/A
APV-I*	40 m	50 m
APV-II*	40 m	20.0 m
Category I precision approach	40 m	35.0 m to 10.0 m

* The terms *APV-I/II* refer to two levels of GNSS *AP*proach and landing operations with *V*ertical guidance.

For detailed definitions, specifications, and explanations, readers are referred to Annex 10 to the Convention on International Civil Aviation ICAO (2023).

These four parameters are valuable for all users and all applications. For example, the User Consultation Platform (UCP), organized periodically by the European Commission and the European Union Agency for the Space Programme (EUSPA), provides the up-to-date GNSS user needs and requirements in the road market segment in its periodically updated report (EUSPA, 2021). Tab. 2.2 outlines the most recent requirements and expectations from the perspective of road and automotive users (based on Annex 7 of EUSPA (2021)).

Table 2.2: Navigation performance parameters for road users.

Application	**Availability**	**Accuracy** (horizontal, vertical)	**Integrity Risk (IR)**	**Coverage**
Safety critical				
(traffic&safety warning)	>0.995	$< 3\,\text{m}/1\,\text{m}$[1], N/A	N/A	Worldwide
(automated driving)	>0.999	$< 20\,\text{cm}$, $< 2\,\text{m}$	10^{-7} to 10^{-8}	Worldwide
Payment critical	>0.995	$< 3\,\text{m}$, N/A	10^{-4}	National
Regulatory critical	>0.995	$< 5\,\text{m}$, N/A	10^{-4}	EU
Smart mobility	>0.995	$< 5\,\text{m}/3\,\text{m}$[2], N/A	N/A	Regional

[1] $< 3\,\text{m}$ for Day 1 applications, and $< 1\,\text{m}$ for advanced applications.
[2] $< 3\,\text{m}$ for scenarios where payment function is included.

2.4 (Advanced) receiver autonomous integrity monitoring

2.4.1 Introduction

Receiver Autonomous Integrity Monitoring (RAIM) is a technique that exploits redundant measurements to achieve self-contained fault detection at the user receiver. The legacy GPS-based RAIM is the most widely used form of satellite navigation by aviation to date (Walter, 2017), e.g., for predicted availability and onboard single-measurement fault monitoring based on single-frequency GPS signals. It is capable of providing horizontal guidance without the need for expensive ground infrastructure and offering global coverage. The main challenge in RAIM is not only designing fault detectors but also evaluating the impact of undetected faults on safety risk (Pullen and Joerger, 2021). RAIM is not designed to detect multiple satellite failures simultaneously.

Advanced RAIM (ARAIM) takes benefits of the GNSS modernization programs, incorporating new frequencies and constellations for aviation use. ARAIM differs from the legacy RAIM in three key areas:

- **Increased number of satellites** from multiple GNSS constellations;
- **Dual-frequency satellite signals** allow the elimination of ionospheric errors;
- **Error and threat model** provided to the airborne receiver in the form of Integrity Support Parameter (ISP) via the Integrity Support Message (ISM).

Moreover, the Protection Level (PL) is a fundamental performance metric for high-integrity GNSS applications. It is computed by users, representing position domain error bounds at the small probabilities required by the integrity requirements, and compared to the Alert

Limit (AL), i.e., the maximum tolerable position errors to maintain safety, that applies to the operation being conducted.

The following sections discuss these elements in detail:

- Sec. 2.4.2 describes the ARAIM threat models and approaches for nominal error bounding, including Gaussian overbounding and nominal bias bounding;
- Sec. 2.4.3 explores fault detection approaches based on test theory: the residual-based (RB) detection and solution separation (SS) detection, utilized in the fault detection and exclusion (FDE) functions in RAIM and ARAIM algorithms;
- Sec. 2.4.4 explains the concepts of loss of integrity (LOI) and loss of continuity (LOC) with corresponding probabilities formulated;
- Sec. 2.4.5 and Sec. 2.4.6 evaluate integrity risk (IR) for classical RAIM and ARAIM, respectively, and introduce respective methods for calculating PL.

2.4.2 Threat modeling and nominal error bounding

Introduction

The ARAIM Technical Subgroup Milestones 1 Report defines navigation threats as all possible events (natural, systemic, or operational) that drive the computed navigation solution to deviate from the true position, regardless of whether a specific fault can be identified in one of the navigation systems or not (WG-C ARAIM TSG, 2012). ARAIM threats are those that impact the performance of the ARAIM algorithm and have a probability of occurrence larger than the required integrity risk (IR).

Eight classes of ARAIM threats have been identified and summarized in Tab. 2.3. A complete table including all identified items can be found in Annex H of the report. Each item refers to a single combination of two threat properties mentioned (nominal/narrow/wide and nature) and GNSS (GPS/Galileo, or cross-constellation).

These threats are categorized into three types:

- **Nominal errors** are errors under nominal conditions when all systems (space, ground, and user segments) operate normally, e.g., receiver noise, multipath, tropospheric delay, inter-frequency bias, nominal signal deformation, nominal orbit determination, and satellite clock errors;
- **Narrow (single) faults** refer to uncorrelated errors affecting satellites individually and not entering into the first category;
- **Wide (multiple) faults** include correlated errors induced by space or ground segment faults that affect navigation signals/messages from multiple satellites and are not included in the first two categories.

The entire threat space must be modeled properly to ensure navigation integrity, including:

- Adequate **nominal error bounding** concerning the first category of threats, which are handled as either
  - **Random errors** that are addressed by Gaussian overbounding methods or
  - **Nominal biases** referring to errors with long-term variations or non-Gaussian behaviors, which are upper bounded by constant values.

Table 2.3: Identified ARAIM threats and their characterizations (WG-C ARAIM TSG, 2012).

Threat source	**Threat category**[1]		
	Nominal	Narrow fault	Wide fault
Clock and ephemeris	Orbit/clock estimation and prediction and broadcast limits	Includes clock runoffs, bad ephemeris, unflagged manoeuvre	Erroneous Earth Orientation Prediction Parameter (EOPP), inadequate manned operations, ground-inherent failures
Signal deformation	Nominal differences in signals due to Radio Frequency (RF) components, filters, and antenna waveform distortion	Failures in satellite payload signal generation components. Faulted signal model as described in ICAO	N/A[3]
Code-carrier incoherence	e.g. incoherence observed in GPS L5 signal or GEO L1 signals[2]		N/A[3]
Inter-frequency bias	Delay differences in satellite payload signal paths[2]		N/A[3]
Satellite antenna bias	Look-angle dependent biases caused at satellite antennas[2]		N/A[3]
Ionospheric	N/A[3]	Scintillation	Multiple scintillations at solar storms in certain latitudes
Tropospheric	Nominal residual troposphere error (after applying SBAS MOPS correction model)	N/A[3]	N/A[3]
Receiver noise and multipath	Nominal noise and multipath terms in airborne model	e.g., receiver tracking failure or multipath from the onboard reflector.	e.g., receiver tracking multiple failures or multipath from the onboard reflector.

[1] ARAIM threats are categorized according to two properties: (i) as a nominal error or fault impacting one or multiple satellites, and (ii) according to the nature of the fault.
[2] The nominal errors and narrow faults for these threat types share the same elaborations in the current version, which may be further distinguished.
[3] Not all possible combinations correspond to real threats, and those deemed irrelevant appear as N/A in the table and should be excluded from any quantitative characterization.

- **Fault probability specification**, capturing the rates of the other two categories of threats.

The subsequent subsections detail the Gaussian overbounding and nominal bias methods and discuss fault probability definitions in ARAIM.

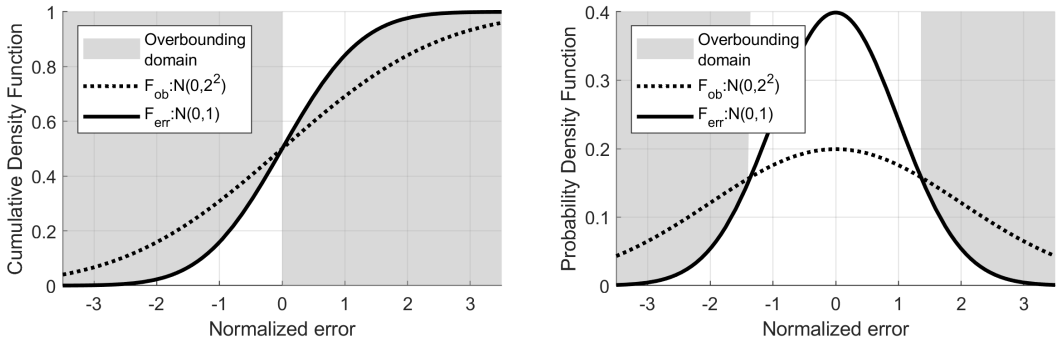
Gaussian overbounding

The normal distribution is often used to model GNSS errors thanks to its simplicity and convenience in combining multiple normally distributed error components. According to the Central Limit Theorem (CLT), as discussed earlier, the sum of independent and identically distributed random variables approaches a normal distribution as the number of variables increases. Therefore, modeling observed errors using a normal distribution with estimated means and variances is generally reasonable for accuracy statistics, e.g., 90% to 95% of the observations. However, the limitations of the CLT challenge this approach at very low probabilities as the actual error distributions have "fatter" tails than normal distributions.

Overbounding is the concept that an actual distribution can be conservatively described by a simple, usually Gaussian, model (DeCleene, 2000; Walter, 2017). The overbounding distribution should predict the probability of large errors to be no lower than the actual distribution. Although the actual error distribution may not be completely known, the overbounding approaches provide a practical way to represent it for analysis. In a mathematical sense, the following requirements are fulfilled:

$$\begin{aligned} F_{ob}(x) &\geq F_{err}(x),\ \forall F_{err}(x) \leq 0.5; \\ F_{ob}(x) &< F_{err}(x),\ \forall F_{err}(x) > 0.5. \end{aligned} \tag{2.23}$$

In Eq. 2.23, $F_{ob}(x)$ is the overbounding Gaussian CDF and $F_{err}(x)$ is the unknown underlying CDF. Subsequently, individual overbounded error components are combined through quadratic variance propagation. Once large errors in the observation domain are overbounded, any subsequent navigation algorithm can satisfy specific mathematical requirements on actual estimation errors (Rife et al., 2006).



(a) CDF domain (b) PDF domain

Figure 2.6: CDF overbounding concept. A standard normal distribution $\mathcal{N}(0,1)$, representing the actual errors (solid curve), can be overbounded by a Gaussian distribution with inflated standard deviation ($\mathcal{N}(0,2^2)$, dotted curve) in the CDF domain. A valid overbounding CDF would always be located within the shadow area, i.e., yielding Eq. 2.23 (left). The tail probabilities of the actual errors are overbounded in the PDF domain (b).

Based on the foundational work of DeCleene (2000), overbounding methods have been expanded in the academic literature to include adjustments in mathematical assumptions such as Rife et al. (2004a) and Blanch et al. (2019b) or by reducing conservatism through alternative distributional models (Rife et al., 2004b; Larson et al., 2019). By trading off

the model complexity, these refined approaches potentially enhance system availability by introducing less conservatism in the range distributions before they are transformed to the position domain (Larson et al., 2019).

In practice, overbounding is often achieved empirically. The sigma value of a Gaussian distribution is varied to fit actual data in terms of CDF. This sigma value is then increased to ensure a sufficient *margin* between the desired overbounding CDF and the actual CDF at the required probabilities. This margin, interpreted as the tightness or conservatism of bounding, typically depends on:

- The shape and tail distribution of the actual data relative to the overbounding model;
- Any theoretical knowledge of the characteristics of the actual distribution.

And it should be determined based on engineering judgment (Pullen and Joerger, 2021).

For instance, Signal-In-Space Range Error (SISRE)[1] represents the equivalent GNSS pseudorange error mainly originating from the satellite clock, broadcast ephemeris, and hardware group delay error sources (i.e., Type 1 threat in Tab. 2.3). Every user r located at a point under the k-th satellite's footprint experiences a distinct Instantaneous User Range Error (IURE), expressed as follows(Perea et al., 2017):

$$\epsilon^k_{IURE,r} = \left(\epsilon^k_{orb}\right)^T \cdot \mathbf{l}^k_r - \epsilon_{clk}, \tag{2.24}$$

where $\mathbf{l}^k_r$ represents the unit vector of the LOS direction; ϵ^k_{orb} and ϵ_{clk} are the instantaneous satellite ephemeris error vector and the clock error term.

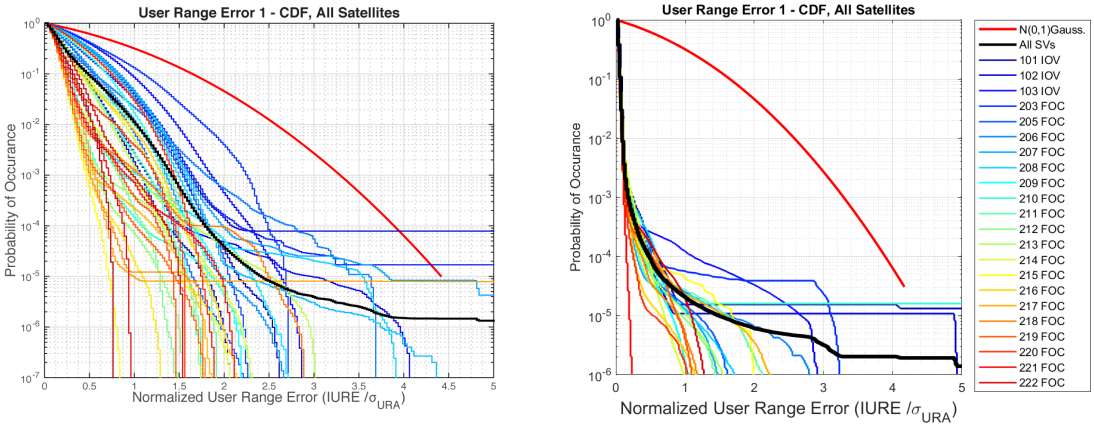
For integrity purposes, Eq. 2.24 is often assessed for its worst-case values, i.e., at a given epoch and for each satellite, the location reporting the greatest IURE is taken into the statistical analysis. Accordingly, Fig. 2.7(a) and 2.7(b) show the historical statistics for GPS and Galileo in the form of one minus Cumulative Density Function (CDF). The thick black curve in both plots represents the aggregate ECDF incorporating all satellites, while the thick red curve illustrates the CDF of an overbounding distribution. The two red lines are extended down to 10^{-5} and 3×10^{-5} probabilities. These probabilities, as will be introduced later, are the *committed* satellite fault probabilities. Nominal Gaussian behavior of SISRE is not required, and not assured, below these levels.

Nominal bias

The nominal bias accounts for near-constant uncorrected errors (such as signal deformation and antenna bias) and non-Gaussian behavior (WG-C ARAIM TSG, 2012). These small-range errors remain essentially constant throughout the duration of the intended operation and, therefore, cannot be treated as purely random. For instance, they are not appropriate to be assumed uncorrelated over periods of time of 15 seconds or more for an aircraft approach. Instead, nominal biases are deterministically upper bounded by a constant value, termed b_{nom}, representing the worst-case situation.

Intensive studies have analyzed nominal range biases. Nominal signal distortions are subtle perturbations on the time domain waveform transmitted by the satellite as observed by the

[1] Other terminologies can be found in literature, such as *Projected (Ranging) Error*, *Signal-In-Space Ranging Error (SISE)*, and *Signal-In-Space User Range Error*, etc. This work chooses the term *Signal-In-Space Range Error (SISRE)* and considers them as synonyms. In addition, it should not be mixed up with the term *User Range Error (URE)*, which denotes an Integrity Support Parameter (ISP), i.e., the standard deviation of SISRE for accuracy and continuity purposes in the ARAIM algorithm.



(a) GPS (Walter and Blanch, 2015) (b) Galileo (Wang and Walter, 2023)

Figure 2.7: The complementary CDF (or 1 minus CDF) of GPS and Galileo Instantaneous User Range Error (IURE), i.e., by individual satellite (colored) and overall satellites (thick black). The GPS values are normalized by the broadcast URA from the GPS legacy navigation data at the time of each measurement (from January 2008 through December 2014), while the normalization of Galileo values is based on the fixed 6-meter value for dual-frequency users (from January 2018 through June 2022). The thick red curves represent the overbounding Gaussian distributions down to the committed probabilities by Constellation Service Provider (CSP).

receiver, resulting from imperfections in the signal generation hardware onboard the satellites. These variations in the transmitted signal cause ranging errors. Additional components include variations in the satellite antenna group delay concerning the pointing angle and user antenna group delay variations with respect to azimuth and elevation (Phelts et al., 2014; Macabiau et al., 2014).

When the AL is set fairly small, these ranging biases cannot be ignored when analyzing the integrity performance. Thus, a term b_{nom} is defined to account for all nominal bias components and broadcast to ARAIM users (WG-C ARAIM TSG, 2012). This term is propagated into the state domain through the least-squares estimator and accommodated in the computation of Protection Level. Macabiau et al. (2015) investigated the impact of nominal biases on the PL computed by the ARAIM baseline algorithm under different settings of the worst-case constant values. This term will be further investigated in Sec. 4.2 and Sec. 6.1.

Fault probabilities

Trustworthy information on the inherent integrity of received GNSS satellite signals is crucial to (Advanced) RAIM. The development of ARAIM in civil aviation led to efforts within the ICAO) Navigation Systems Panel (NSP) to obtain additional integrity-related commitments from Constellation Service Provider (CSP). The up-to-date Integrity Support Parameter (ISP) are provided in Tab. 2.4, with the following terms:

- **User Range Accuracy (URA)** is a standard deviation value of a zero-mean Gaussian distribution, which is chosen such that actual SISRE can be upper bounded to the probability specified by P_{sat}, as in Fig. 2.7. From this perspective, the overbounding Gaussian distribution for the nominal SISRE can be described by $\mathcal{N}(0, \text{URA}^2)$, whereas larger errors with probabilities below P_{sat} are treated as anomalous (or faulted) and are not described by this distribution. URA is a more conservative concept than the similar term **User Range Error (URE)**, which also expresses typical range errors at the $1 \cdot \sigma$ level but is not meant to bound errors out to P_{sat} (i.e., URA > URE). The Galileo Open

Table 2.4: Constellation ISP commitments.

Parameter	GNSS constellation			
	GPS	GLONASS	Galileo	Beidou
URA [m]	in NAV data (ICD table)	18	6 (dual frequency), 7.5 (single frequency)	in NAV data (ICD table)
URA Threshold[1] [m]	4.42×URA	70 (3.89×URA)	25 (dual frequency), 31.3 (single frequency)	4.42×URA
P_{sat}	10^{-5}	10^{-4}	3×10^{-5}	10^{-5}
P_{const}	10^{-8}	10^{-4}	2×10^{-4}	6×10^{-5}
TTA [sec]	10	10	not specified	10

Service (OS) adopts the term **Signal-In-Space Accuracy (SISA)**, nevertheless, with the same concept.

- **URA Threshold** defines the SISRE level beyond which errors are considered anomalous or faulted as opposed to nominal (at the same time with probabilities below P_{sat}). It is typically in the form of a multiple of URA. The URA thresholds for Galileo and GLONASS are fixed values since they offer constant URA as in Tab. 2.4, whereas the one for GPS varies with its broadcast URA values.
- **Satellite fault probability (P_{sat})** is a conservative bound on the per-satellite failure state probability for statistically independent satellite failures. It represents the probability at any given time that a satellite will be in a faulty (as opposed to nominal) state. It is an input to the baseline ARAIM algorithm developed for civil aviation and can be determined based on statistical analysis against the satellite failure probabilities.
- **Constellation fault probability (P_{const})** is a conservative bound on the state probability of failures correlated across satellites in the same constellation, including faults of two or more up to all satellites in the same constellation. Such correlated faults affect the entire constellation and can only be detected by checking the consistency with other (non-faulted) constellations.
- **Time-to-Alert (TTA)** denotes the maximum interval between the time at which a faulted condition begins (based on the aforementioned definitions) and when it must end or be alerted to prevent it from being considered a fault from the point of view of potential loss of integrity.

When only satellite-wise faults are considered, Joerger et al. (2012) computes the probability for the fault modes with n_f faulty signals using the binomial law:

$$\begin{aligned} P(n_f = 1) &= \binom{n}{1} P_{sat} \cdot (1 - P_{sat})^{n-1}, \\ P(n_f = 2) &= \binom{n}{2} P_{sat}^2 \cdot (1 - P_{sat})^{n-2}, \\ P(n_f \geq 3) &= \sum_{k=3}^{n} \binom{n}{k} P_{sat}^k \cdot (1 - P_{sat})^{n-k}. \end{aligned} \tag{2.25}$$

with n denoting the total number of satellites.

The baseline ARAIM algorithm in WG-C ARAIM TSG (2022) uses P_{event} to represent an independent fault event, which can be a single satellite failure or a constellation failure:

$$\begin{aligned} &P_{event\ 1} = P_{sat\ 1},\ ...,\ P_{event\ n} = P_{sat\ n}, \\ &P_{event\ n+1} = P_{const\ 1},\ ...,\ P_{event\ n+n_c} = P_{const\ n_c}, \end{aligned} \tag{2.26}$$

providing n_c constellations, and hence, a total of $n+n_c$ events are observed. For a given fault mode, as an example, referring to a fault in satellite 1 and no fault in any constellation, its probability is computed using the binomial law:

$$P(\text{Fault mode 1}) = P_{sat\ 1} \cdot \prod_{i=2}^{n} (1 - P_{sat\ i}) \cdot \prod_{j=1}^{n_c} (1 - P_{const\ j}) \tag{2.27}$$

The general form for the probability of fault mode k is

$$\begin{aligned} P(\text{Fault mode } k) &= \prod_{i=1}^{n+n_c} P_{event\ i}^{w_{i,k}} \cdot (1 - P_{event\ i})^{1-w_{i,k}} \\ \text{with } w_{i,k} &= \begin{cases} 1, & \text{if event i is in fault mode } k; \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{2.28}$$

Notably, it is unnecessary to compute all fault modes for computational efficiency since some fault modes with multiple simultaneous faults may be fairly rare compared to related risk requirements. Reducing the number of monitored fault modes has drawn significant attention in developing ARAIM.

2.4.3 Test theory-based fault detection

Measurement model

The following measurement model is studied:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e} + \mathbf{f}, \tag{2.29}$$

where,

- $\mathbf{y}$ is an $n \times 1$ vector of measurements;
- $\mathbf{A}$ is an $n \times m$ design matrix, indicating the Line-of-Sight (LOS) vectors;
- $\mathbf{x}$ is an $m \times 1$ vector of states, e.g., parameters to be estimated;
- $\mathbf{e}$ is an $n \times 1$ vector of nominal measurement errors;
- $\mathbf{f}$ is an $n \times 1$ vector of faults.

It is assumed that the nominal measurement errors are bounded by a zero-mean multivariate Gaussian distribution, i.e., $\mathbf{e} \sim \mathcal{N}(\mathbf{0}_{n\times 1}, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is a known nonsingular covariance matrix with variances for each measurement on the diagonal, i.e., $\sigma_1^2, ..., \sigma_n^2$.

By manipulating the matrix operation $\mathbf{\Sigma} = (\mathbf{\Sigma}^{\frac{1}{2}})^T \mathbf{\Sigma}^{\frac{1}{2}}$ and multiplying $\mathbf{\Sigma}^{-\frac{1}{2}}$ with both side of Eq. 2.29, the measurement model can be normalized:

$$\begin{aligned} \tilde{\mathbf{y}} =& \tilde{\mathbf{A}}\,\mathbf{x} + \tilde{\mathbf{e}} + \tilde{\mathbf{f}}, \\ \text{with } \tilde{\mathbf{y}} =& \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{y}, \\ \tilde{\mathbf{A}} =& \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{A}, \\ \tilde{\mathbf{e}} =& \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n), \\ \tilde{\mathbf{f}} =& \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{f}. \end{aligned} \tag{2.30}$$

Least-squares estimation

The (weighted) least-squares estimator is used in classical GNSS integrity monitoring methods based on certain stochastic assumptions. Section 4.2 revisits the least-squares estimator to explore corresponding error bounding approaches in a broader context.

Considering the measurement model described in Eq. 2.29, the least-squares estimate of the state, denoted by $\hat{\mathbf{x}}$, is expressed as follows:

$$\hat{\mathbf{x}} = \mathbf{K}\,\mathbf{y}, \tag{2.31}$$

where $\mathbf{K} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{P}$ with $\mathbf{P} = \mathbf{\Sigma}^{-1}$. The least-squares estimate of measurement reads:

$$\hat{\mathbf{y}} = \mathbf{A}\mathbf{K}\,\mathbf{y}. \tag{2.32}$$

Notably, for the normalized measurement model in Eq. 2.30, the same least-squares estimates of the states can be derived with the counterpart $\tilde{\mathbf{K}}$ matrix:

$$\tilde{\mathbf{K}} = \left(\tilde{\mathbf{A}}^T\tilde{\mathbf{A}}\right)^{-1}\tilde{\mathbf{A}}^T = \left(\mathbf{A}^T\left(\mathbf{P}^{\frac{1}{2}}\right)^T\mathbf{P}^{\frac{1}{2}}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{P}^{\frac{1}{2}}, \tag{2.33}$$

where $\mathbf{P}^{\frac{1}{2}} = \mathbf{\Sigma}^{-\frac{1}{2}}$, and the counterpart $\hat{\tilde{\mathbf{y}}}$, the least-squares estimate of normalized measurements, is expressed as

$$\begin{aligned}\hat{\tilde{\mathbf{y}}} = \tilde{\mathbf{A}}\tilde{\mathbf{K}}\tilde{\mathbf{y}} =& \mathbf{P}^{\frac{1}{2}}\mathbf{A}\left(\mathbf{A}^T\left(\mathbf{P}^{\frac{1}{2}}\right)^T\mathbf{P}^{\frac{1}{2}}\mathbf{A}\right)^{-1}\mathbf{A}^T\left(\mathbf{P}^{\frac{1}{2}}\right)^T\mathbf{P}^{\frac{1}{2}}\mathbf{y}\\ =& \mathbf{P}^{\frac{1}{2}}\mathbf{A}\mathbf{K}\,\mathbf{y}\end{aligned} \tag{2.34}$$

An $m \times 1$ vector α is defined to extract the state of interest, e.g., the vertical position for aircraft precision approach navigation using GPS signals:

$$\alpha^T = \begin{bmatrix}0 & 0 & 1 & 0\end{bmatrix}, \tag{2.35}$$

where only one element of α is one, corresponding to the state of interest based on the order in which states are stacked in $\mathbf{x}$, and the other $m-1$ elements are zeros. As a result, the m-th state to be estimated is

$$\hat{x}_{(m)} = \mathbf{k}_{(m)}^T\mathbf{y}, \tag{2.36}$$

where $\mathbf{k}_{(m)}^T = \alpha^T\mathbf{K}$.

The estimation error is

$$\varepsilon_{(m)} \coloneqq \hat{x}_{(m)} - x_{(m)} = \mathbf{k}_{(m)}^T(\mathbf{e} + \mathbf{f}), \tag{2.37}$$

which follows a Gaussian distribution according to the law of quadratic error propagation:

$$\varepsilon_{(m)} \sim \mathcal{N}\left(\mathbf{k}_{(m)}^T\mathbf{f},\, \sigma_{(m)}^2 = \alpha^T\mathbf{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}\alpha\right), \tag{2.38}$$

with the Variance-Covariance Matrix (VCM) of state estimates $\mathbf{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}$.

Residual-based method

The residual vector is defined as (Parkinson and Axelrad, 1988; Potter and Suman, 1977)

$$\begin{aligned}\mathbf{r} \coloneqq \mathbf{y} - \hat{\mathbf{y}} =& (\mathbf{I}_n - \mathbf{AK})\mathbf{y} = (\mathbf{I}_n - \mathbf{A}(\mathbf{A}^T\mathbf{PA})^{-1}\mathbf{A}^T\mathbf{P})(\mathbf{Ax} + \mathbf{e} + \mathbf{f}) \\ =& (\mathbf{I}_n - \mathbf{AK})(\mathbf{e} + \mathbf{f}).\end{aligned} \tag{2.39}$$

A normalized residual vector $\tilde{\mathbf{r}}$ for Eq. 2.30 is defined as

$$\tilde{\mathbf{r}} \coloneqq \tilde{\mathbf{y}} - \hat{\tilde{\mathbf{y}}} = \mathbf{P}^{\frac{1}{2}}\mathbf{y} - \mathbf{P}^{\frac{1}{2}}\mathbf{AKy} = \mathbf{P}^{\frac{1}{2}}\mathbf{r} \tag{2.40}$$

The residual-based (RB) test statistic q_{RB}^2 is the weighted norm of the residuals:

$$q_{RB}^2 \coloneqq \mathbf{r}^T\mathbf{Pr} = \tilde{\mathbf{r}}^T\tilde{\mathbf{r}}. \tag{2.41}$$

The scalar random variable q_{RB}^2 follows a non-central chi-square distribution with $(n-m)$ degrees of freedom and non-centrality parameter λ_{RB} in the presence of faults:

$$\begin{gathered}q_{RB}^2 \sim \chi^2(n-m, \lambda_{RB}), \\ \text{with } \lambda_{RB} = \mathbf{f}^T\mathbf{Sf} \text{ and } \mathbf{S} = \mathbf{P}(\mathbf{I}_n - \mathbf{AK}).\end{gathered} \tag{2.42}$$

Additionally, the RB test statistic can be derived from the parity vector $\mathbf{p}$, which lies in the $(n-m)$-dimensional parity space, or left null space of $\tilde{\mathbf{A}}$:

$$\mathbf{p} \coloneqq \mathbf{Q}\tilde{\mathbf{y}} = \mathbf{Q}(\tilde{\mathbf{e}} + \tilde{\mathbf{f}}), \tag{2.43}$$

where $\mathbf{Q}$ is the $(n-m) \times n$ parity matrix, defined by two conditions:

$$\begin{aligned}&\mathbf{QQ}^T = \mathbf{I}_{n-m}, \text{ and} \\ &\mathbf{Q}\tilde{\mathbf{A}} = \mathbf{0}_{(n-m)\times m}.\end{aligned} \tag{2.44}$$

The normalized measurement vector $\tilde{\mathbf{y}}$ can be broken down into two orthogonal components:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{A}}\tilde{\mathbf{K}}\tilde{\mathbf{y}} + \mathbf{Q}^T\mathbf{Q}\tilde{\mathbf{y}}, \tag{2.45}$$

with the first component $\tilde{\mathbf{AK}}\tilde{\mathbf{y}}$ lying in the column space of $\tilde{\mathbf{A}}$, and the second component $\mathbf{Q}^T\mathbf{Q}\tilde{\mathbf{y}}$ in the null space of $\tilde{\mathbf{A}}^T$, respectively. Subsequently, one gets

$$\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n - \tilde{\mathbf{A}}\tilde{\mathbf{K}} = \mathbf{I}_n - \mathbf{P}^{\frac{1}{2}}\mathbf{A}(\mathbf{A}^T\mathbf{PA})^{-1}\mathbf{A}^T\mathbf{P}^{\frac{1}{2}}. \tag{2.46}$$

Substituting the results in Eq. 2.46 into Eq. 2.39 and Eq. 2.41 shows that the L_2-norms of parity vector $\mathbf{p}$ and of the normalized residual vector $\tilde{\mathbf{r}}$ are equal:

$$q_{RB}^2 = \tilde{\mathbf{r}}^T\tilde{\mathbf{r}} = \tilde{\mathbf{y}}^T(\mathbf{I}_n - \tilde{\mathbf{A}}\tilde{\mathbf{K}})^T(\mathbf{I}_n - \tilde{\mathbf{A}}\tilde{\mathbf{K}})\tilde{\mathbf{y}} = \mathbf{p}^T\mathbf{p}. \tag{2.47}$$

Multiple Hypotheses Solution Separation (MHSS)

The MHSS framework accounts for the entire set of mutually exclusive, jointly exhaustive hypotheses, including the null hypothesis (H_0, the fault-free situation) and alternative hypotheses ($H_i,\ i = 1, ..., h$, the faulty situations). Each alternative hypothesis H_i indicates a

fault mode where a subset of measurements is faulty, while the rest are fault-free. Accordingly, the measurement model in Eq. 2.29 is partitioned to distinguish the subset of faulty measurements $\mathbf{H}_i^{*T}\mathbf{y}$ from the fault-free measurements $\mathbf{H}_i^T\mathbf{y}$:

$$\begin{aligned} \begin{bmatrix} \mathbf{H}_i^T\mathbf{y} \\ \mathbf{H}_i^{*T}\mathbf{y} \end{bmatrix} &= \begin{bmatrix} \mathbf{H}_i^T\mathbf{A} \\ \mathbf{H}_i^{*T}\mathbf{A} \end{bmatrix}\mathbf{x} + \begin{bmatrix} \mathbf{H}_i^T(\mathbf{e}+\mathbf{f}) \\ \mathbf{H}_i^{*T}(\mathbf{e}+\mathbf{f}) \end{bmatrix}, \\ \text{with } \mathbf{H}_i &\coloneqq \begin{bmatrix} \mathbf{0}_{n_i\times(n-n_i)} \\ \mathbf{I}_{n-n_i} \end{bmatrix} \text{ and } \mathbf{H}_i^* \coloneqq \begin{bmatrix} \mathbf{I}_{n_i} \\ \mathbf{0}_{(n-n_i)\times n_i} \end{bmatrix}. \end{aligned} \tag{2.48}$$

In Eq. 2.48, the matrices $\mathbf{H}_i$ and $\mathbf{H}_i^*$ are used for extraction. Subsequently, two new geometry matrices can be constructed from $\mathbf{A}$ for all alternative hypotheses:

$$\begin{bmatrix} \mathbf{A}_i \\ \mathbf{A}_i^* \end{bmatrix} \coloneqq \begin{bmatrix} \mathbf{H}_i^T\mathbf{A} \\ \mathbf{H}_i^{*T}\mathbf{A} \end{bmatrix}. \tag{2.49}$$

If the number of fault measurements in H_i is denoted by n_i, $\mathbf{A}_i$ contains the rows corresponding to n_i fault-free measurements and the other rows of zeros. In contrast, $\mathbf{A}_i^*$ comprises the non-zero rows corresponding to $n - n_i$ faulty measurements.

Without loss of generality, applying this rule to the null hypothesis H_0 leads to $\mathbf{A}_0 = \mathbf{A}$ and $\mathbf{A}_0^*$ being a $n \times m$ matrix of zeros ($n_0 = 0$).

Under H_i, the least-squares subset solution $\hat{x}_{i,(m)}$ for the m-th state of interest using $n - n_i$ fault-free measurements is:

$$\begin{aligned} \hat{x}_{i,(m)} &\coloneqq \mathbf{k}_{i,(m)}^T\mathbf{y}, \text{ for } i = 1, ..., h, \\ \text{with } \mathbf{k}_{i,(m)}^T &= \alpha\,\mathbf{K}_i, \\ \mathbf{K}_i &= (\mathbf{A}_i^T\mathbf{P}_i\mathbf{A}_i)^{-1}\mathbf{A}_i^T\mathbf{P}_i\mathbf{H}_i^T = (\mathbf{A}^T\mathbf{H}_i\mathbf{H}_i^T\mathbf{P}\mathbf{H}_i\mathbf{H}_i^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{H}_i\mathbf{H}_i^T\mathbf{P}\mathbf{H}_i\mathbf{H}_i^T. \end{aligned} \tag{2.50}$$

As SS detection always performs on a single state, this section omits the subscript "(m)" for simplicity.

It is worth clarifying that under H_0, $\hat{x}_0$ refers to the fault-free solution derived using the full set of measurements. In contrast, under H_i ($i \neq 0$), $\hat{x}_0$ represents the full-set solution, which in this case is not fault-free. The remainder of this section uses numbered subscripts i ranging from 0 to h, unless stated otherwise.

Subsequently, the estimation error ε_i under H_i is given by

$$\begin{aligned} \varepsilon_i &\coloneqq \mathbf{k}_i^T(\mathbf{e}+\mathbf{f}), \\ \text{and } \varepsilon_i &\sim \mathcal{N}\left(0, \sigma_i^2 = \alpha^T\mathbf{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}},i}\alpha\right), \end{aligned} \tag{2.51}$$

with the VCM $\mathbf{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}},i} = (\mathbf{A}_i^T\mathbf{P}\mathbf{A}_i)^{-1}$.

The solution separation ($\Delta_{SS,i}$ for $i = 1, ..., h$) is defined as (Lee, 1986; Brenner, 1995)

$$\Delta_{SS,i} \coloneqq \hat{x}_0 - \hat{x}_i = \varepsilon_0 - \varepsilon_i. \tag{2.52}$$

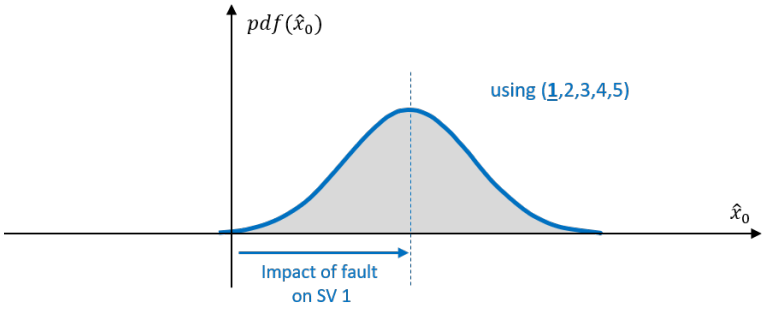
$\Delta_{SS,i}$ can also be expressed as

$$\begin{aligned} \Delta_{SS,i} &= (\mathbf{k}_0 - \mathbf{k}_i)^T\mathbf{y} = (\mathbf{k}_0 - \mathbf{k}_i)^T(\mathbf{e}+\mathbf{f}), \\ \text{with } \Delta_{SS,i} &\sim \mathcal{N}\left((\mathbf{k}_0 - \mathbf{k}_i)^T\mathbf{f}, \sigma_{\Delta,i}^2\right), \\ \sigma_{\Delta,i}^2 &= \alpha^T\left(\mathbf{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}},i} - \mathbf{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}},\mathbf{0}}\right)\alpha = \sigma_i^2 - \sigma_0^2. \end{aligned} \tag{2.53}$$

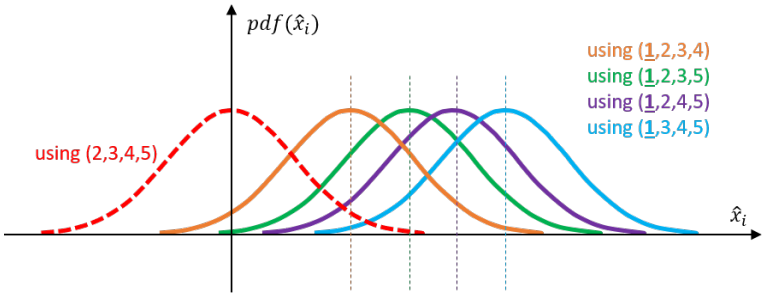
Similar to the RB test statistic, a *normalized* SS test statistic is defined as

$$\begin{aligned} q_{SS,i} &:= \frac{\Delta_{SS,i}}{\sigma_{\Delta,i}} \text{ for } i = 1, ..., h, \\ \text{and } q^2_{SS,i} &\sim \chi^2(1, \lambda_i) \text{ with } \lambda_i = \frac{\mathbf{f}^T(\mathbf{k}_0 - \mathbf{k}_i)(\mathbf{k}_0 - \mathbf{k}_i)^T\mathbf{f}}{\sigma^2_{\Delta,i}}. \end{aligned} \tag{2.54}$$

Fig. 2.8 illustrates the principle of the SS test statistic.



(a) 1 full-set solution using 5 satellites



(b) 5 subset solutions using 4 satellites

Figure 2.8: Concepts of solution separation detection. With five satellites available, six combinations of at least four satellites can be formed for estimation. Fig. 2.8(a) illustrates the full-set solution for a single state, which is biased by a fault on Satellite 1. In Fig. 2.8(b) of subset solutions, the solutions involving Satellite 1 all have their mean shifted with respect to the fault-free subset (using Satellite 2-5, red dashed curve). The maximum solution separation $\max_{i\in\{1,...,5\}} \Delta_{SS,i} = \max_{i\in\{1,...,5\}} (\hat{x}_0 - \hat{x}_i)$ is an intuitive and efficient detection test statistic (Pullen and Joerger, 2021).

The normalized SS test statistic is a projection of the parity vector on the corresponding fault mode line (Joerger et al., 2014):

$$\begin{aligned} q_{SS,i} &= \frac{(\mathbf{k}_0 - \mathbf{k}_i)^T}{\sigma_{\Delta,i}}\mathbf{y} = \mathbf{u}_i^T\mathbf{p}, \\ \mathbf{u}_i^T &:= \frac{\mathbf{Q}\mathbf{H}_i^*}{\sqrt{\mathbf{H}_i^{*T}\mathbf{Q}^T\mathbf{Q}\mathbf{H}_i^*}}, \text{ for } i = 1, ..., n, \end{aligned} \tag{2.55}$$

where $\mathbf{Q}$ is the matrix defined in Eq. 2.44. The vector $\mathbf{u}_i$ indicates the direction of the fault line passing through the origin in parity space, corresponding to a fault on the i-th measurement in the presence of single-measurement faults.

For further theoretical comparison between the two methods, the readers are referred to literature, e.g., Joerger and Pervan (2014); Joerger et al. (2014); Joerger and Pervan (2016).

2.4.4 Loss of integrity and loss of continuity

Loss of continuity

The continuity risk (CR), or probability of loss of continuity (LOC) $P(\text{LOC})$, is the probability of a detected but unscheduled navigation function interruption after an operation has been initiated (Zhai et al., 2015) for aviation applications. It includes:

- the probability due to false alert (FA), i.e., $P(\text{FA})$,
- the probability due to fault detection when a fault occurs, and
- the probability due to other causes of loss of continuity.

In practice, the FA probability is limited by an allocated CR requirement $C_{REQ,0}$. The other components are typically evaluated dependent on dedicated fault models, which is beyond the scope of this work. Relevant information can be found in, e.g., Zhai et al. (2015), Joerger and Pervan (2016), and Zhai (2018).

For the RB detector, the CR is dependent on the test threshold T_{RB}:

$$P\left(|q_{RB}| \geq T_{RB} \mid H_0\right) \cdot P(H_0) \leq C_{REQ,0}. \tag{2.56}$$

T_{RB} can be determined once $C_{REQ,0}$ is given using the following equation:

$$P\left(|q_{RB}| \geq T_{RB} \mid H_0\right) = \int_{T_{RB}^2}^{+\infty} \chi^2_{q_{RB}}(n-m, 0)\ \mathrm{d}q_{RB}^2. \tag{2.57}$$

The SS detection thresholds $T_{SS,i}$ for each fault mode i are associated with $C_{REQ,0}$:

$$P\left(|q_{SS,1}| \geq T_{SS,1} \vee \ldots \vee |q_{SS,h}| \geq T_{SS,h} \mid H_0\right) \cdot P(H_0) \leq C_{REQ,0}. \tag{2.58}$$

The CR budget for SS detection is influenced by all fault modes and thus should account for multiple test statistics to ensure that the overall requirement is satisfied (Joerger and Pervan, 2014). A precise assessment can be done on the joint probability based on multidimensional PDFs. An upper bounding approach is usually employed for real-time implementations:

$$\begin{aligned} &P\left(|q_{SS,1}| \geq T_{SS,1} \vee \ldots \vee |q_{SS,h}| \geq T_{SS,h} \mid H_0\right) \cdot P(H_0) \\ \leq &\sum_{i=1}^{h} P\left(|q_{SS,i}| \geq T_{SS,i} \mid H_0\right) \cdot P(H_0). \end{aligned} \tag{2.59}$$

Subsequently, the SS detection thresholds $T_{SS,i}$ can be computed by

$$\begin{aligned} T_{SS,i} &= Q^{-1}\left\{\frac{C_{REQ,i}}{2 \cdot P(H_0)}\right\}, \\ C_{REQ,0} &= \sum_{i=1}^{h} C_{REQ,i},\ \text{e.g., } C_{REQ,i} = \frac{C_{REQ,0}}{h}, \end{aligned} \tag{2.60}$$

where $Q^{-1}(P)$ denotes the $(1-P)$ quantile of a standard normal distribution, i.e., $Q(x) = 1 - F_\Phi(x)$ where $F_\Phi(\cdot)$ is the CDF of a standard normal distribution. The CR can be allocated arbitrarily, e.g., the equal allocation is used in Eq. 2.60. This ensures that the CR requirement is met but may cause the integrity risk bound to be loose.

Loss of integrity

The integrity risk (IR), or equivalently the probability of Hazardous Misleading Information (HMI) or loss of integrity (LOI), is a joint probability:

$$P(\text{HMI}) = P(|\varepsilon_0| > \ell \wedge |q| < T), \tag{2.61}$$

with

- ε_0 the estimation error using full-set measurements,
- ℓ a specific Alert Limit (AL) that defines the hazardous situations,
- q RB/SS detection test statistic (q_{RB}/q_{SS}),
- T RB/SS detection thresholds (T_{RB}/T_{SS}),

which involves the event of *Position Failure* (PF, or termed *Hazardous Information* (HI), $|\varepsilon_0| > \ell$), and the *No Detection* (ND) event ($|q| < T$).

Given an alternative hypothesis H_i, the HMI probability is set to

$$P(\text{HMI} \mid H_i) = P(|\varepsilon_0| > \ell \wedge |q| < T \mid H_i). \tag{2.62}$$

For the multiple-hypothesis approach, all hypotheses and test statistics must be considered when evaluating the HMI probability:

$$P(\text{HMI}) = \sum_{i=0}^{h} P(|\varepsilon_0| > \ell \wedge |q_{SS,1}| < T_{SS,1} \wedge ... \wedge |q_{SS,h}| < T_{SS,h} \mid H_i) \cdot P(H_i) \tag{2.63}$$

Here, $P(H_i)$ is interpreted as the probability of a specific fault mode, discussed in Sec. 2.4.2.

2.4.5 Classical RAIM

Integrity risk evaluation

The classical residual-based RAIM is designed for GPS-based navigation against potential single-measurement faults. Binary hypotheses typically considered are as follows:

- H_0: fault-free hypothesis.
- H_A: faulty hypothesis.

The conditional HMI probability under H_0 can be upper bounded by

$$\begin{aligned} P(\text{HMI} \mid H_0) &= P(|\varepsilon_0| > \ell \wedge |q_{RB}| < T_{RB} \mid H_0) \\ &\leq P(|\varepsilon_0| > \ell \mid H_0), \end{aligned} \tag{2.64}$$

where the upper bound $P(|q_{RB}| < T_{RB} \mid H_0) = 1 - P(\text{FA}) \leq 1$ is used since the FA probability is limited by the CR requirement.

A bound on the HMI probability given H_A can be evaluated for the $n \times 1$ worst-case fault (WCF) vector $\mathbf{f}_{\text{worst}}$:

$$P(|\varepsilon_0| > \ell \wedge |q_{RB}| < T_{RB} \mid H_A) \leq P(|\varepsilon_0| > \ell \wedge |q_{RB}| < T_{RB} \mid \mathbf{f}_{\text{worst}}). \tag{2.65}$$

$\mathbf{f}_{\text{worst}}$ is defined such that the IR is maximized for the alternative hypothesis H_A.

It is shown in literature, e.g., Potter and Suman (1977), that the least-squares estimate error is independent of the RB test statistic q_{RB}. Subsequently, Joerger et al. (2012) rewrites the joint probability in Eq. 2.62 as a product of probabilities:

$$\begin{aligned} P(\text{HMI} \mid H_A) &= P(|\varepsilon_0| > \ell \mid H_A) \cdot P(|q_{RB}| < T_{RB} \mid H_A) \\ &\leq P(|\varepsilon_0| > \ell \mid \mathbf{f}_{\text{worst}}) \cdot P(|q_{RB}| < T_{RB} \mid \mathbf{f}_{\text{worst}}). \end{aligned} \tag{2.66}$$

The direction of $\mathbf{f}_{\text{worst}}$ can be found utilizing the knowledge of geometry; for example, Joerger and Pervan (2014) provides analytical solutions. Subsequently, the magnitude of $\mathbf{f}_{\text{worst}}$ can be determined using a line search method.

The ND probability $P(|q_{RB}| < T_{RB} \mid \mathbf{f}_{\text{worst}})$ can be evaluated considering the RB test statistic's distributions (cf Eq. 2.42). The PF probability $P(|\varepsilon_0| > \ell \mid \mathbf{f}_{\text{worst}})$ is evaluated based on the estimation error distribution cf. Eq. 2.37 and Eq. 2.38.

Furthermore, Joerger and Pervan (2014) establishes a tight IR bound for RB RAIM in the multiple hypotheses framework:

$$\begin{aligned} P(\text{HMI}) =& P(\text{HMI} \mid H_0) \cdot P(H_0) + \sum_{i=1}^{h} P(\text{HMI} \mid H_i) \cdot P(H_i) \\ \leq& P(|\varepsilon_0| > \ell \mid H_0) \cdot P(H_0) + \\ & \sum_{i=1}^{h} P(|\varepsilon_0| > \ell \mid \mathbf{f}_{i,\text{worst}}) \cdot P(|q_{RB}| < T_{RB} \mid \mathbf{f}_{i,\text{worst}}) \cdot P(H_i), \end{aligned} \tag{2.67}$$

where $\mathbf{f}_{i,\text{worst}}$ represents the worst-case fault vector in the i-th fault mode. It should be mentioned that Eq. 2.67 is only for evaluation purposes and not for real-world implementations.

PL computation

In classical RAIM, the PL is computed as a conservative upper bound in the state domain, capturing the "worst" satellite geometry. Various approaches have been developed to address this issue, such as Walter and Enge (1995), Brown and Chin (1997), Milner and Ochieng (2011), etc., and to extend to multiple-fault situations in Angus (2006) and Liu et al. (2022). This section aims to introduce a representative method from Brown and Chin (1997), which will be implemented in the experimental evaluation for comparison.

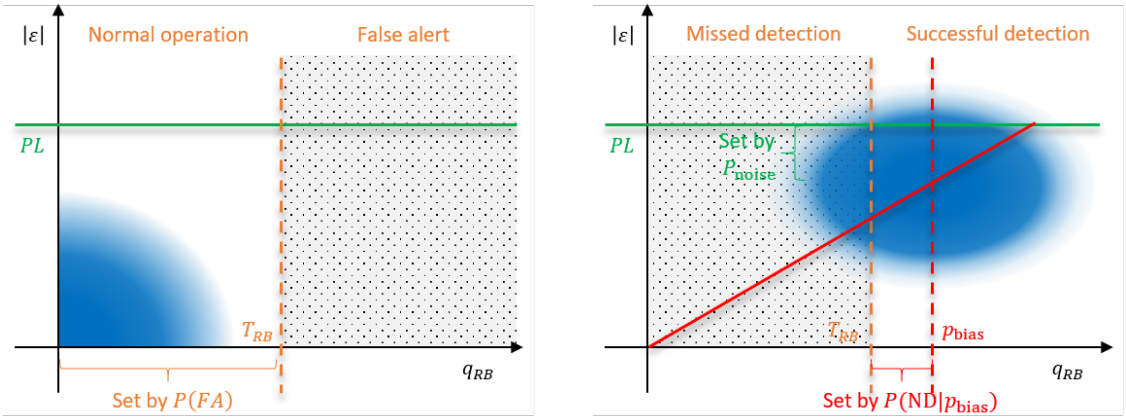
Brown and Chin (1997) takes advantage of the SLOPE concept that characterizes the relationships between the range biases (fault) in each measurement and the induced position errors. For example, the detection threshold is determined based on a required FA probability in Fig. 2.9(a); the red solid line in Fig. 2.9(b) illustrates the relationship between the test statistic (q_{RB}) and the position error ($|\varepsilon|$) under the no-noise assumption. The slope of the line is a function of the geometry matrix and indicates the difficulty of detecting a bias (fault) in the respective measurement. The slope for horizontal positioning associated with the i-th measurement is

$$\text{HSLOPE}_i = \frac{\sqrt{\mathbf{K}_{1i}^2 + \mathbf{K}_{2i}^2}}{\sqrt{\mathbf{S}_{ii}}}, \tag{2.68}$$

where $\mathbf{S}$ is defined in Eq. 2.42, and the indices indicate the particular elements of a matrix.

Eq. 2.68 can be understood as follows. Given a fault vector $\mathbf{f} = [0, ..., 0, f, 0, ..., 0]^T$, i.e., the i-th measurement is faulty with the fault magnitude f, the noise-free assumption would produce a position bias on the horizontal plane based on Eq. 2.37 and Eq. 2.38:

$$\left\| \begin{bmatrix} E(\varepsilon_{(1)}) \\ E(\varepsilon_{(2)}) \end{bmatrix} \right\| = \left\| \left[\mathbf{k}_{(1)}, \mathbf{k}_{(2)} \right]^T \cdot \mathbf{f} \right\| = \sqrt{(\mathbf{K}_{1i})^2 + (\mathbf{K}_{2i})^2} \cdot f. \tag{2.69}$$



(a) Fault-free hypothesis (H_0)

(b) Alternative hypothesis (H_A)

Figure 2.9: Conceptual sketch for the distribution of position errors and the RB test statistic for both fault-free and in the case of a faulty measurement (Brown and Chin, 1997).

According to Eq. 2.41, the induced bias in the test statistic q_{RB} is represented by the square root of the non-centrality parameter in Eq. 2.42:

$$\sqrt{\lambda_{RB}} = \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = \sqrt{\mathbf{S}_{ii}} \cdot f. \tag{2.70}$$

The slope of the fault line associated with the i-th measurement (Eq. 2.68) can be determined by the ratio of Eq. 2.69 and Eq. 2.70.

Fig. 2.9(b) provides another interpretation. The projection of the cloud's center onto the vertical axis can be determined once its projection on the horizontal axis is known.

It is claimed that a fault in the direction associated with the largest slope is the most difficult to detect and yields the smallest test statistic for a given position error:

$$\text{HSLOPE}_{\max} = \max_{i=1,2,\ldots,n} \left\{ \frac{\mathbf{K}_{1i}^2 + \mathbf{K}_{2i}^2}{\sqrt{\mathbf{S}_{ii}}} \right\}. \tag{2.71}$$

Subsequently, the HPL is computed as a sum of two terms:

- The first term protects against biases (undetected faults), involving the maximum slope $\text{HSLOPE}_{\max}$ and a deterministic bias p_{bias}. p_{bias} is defined such that the ND probability conditioned on $\sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = p_{bias}$

  $$P(|q_{RB}| < T_{RB} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = p_{\text{bias}}) = \int_0^{T_{RB}^2} \chi^2_{q_{RB}}(n-m, p^2_{\text{bias}}) \tag{2.72}$$

  is limited by the IR requirement (I_{REQ}). The detection threshold T_{RB} has been determined using Eq. 2.57. It can be seen that any faults yielding $\sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} \geq p_{\text{bias}}$ would satisfy

  $$\begin{aligned} &P(|q_{RB}| < T_{RB} \wedge |\varepsilon_0| > \text{HSLOPE}_{\max} \cdot p_{bias} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} \geq p_{\text{bias}}) \\ =&P(|q_{RB}| < T_{RB} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} \geq p_{\text{bias}}) \cdot P(|\varepsilon_0| > \text{HSLOPE}_{\max} \cdot p_{bias} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} \geq p_{\text{bias}}) \\ \leq&I_{REQ}. \end{aligned} \tag{2.73}$$

  Eq. 2.73 implies that the first term alone can protect against relatively large faults. Specially, it has an upper bound of $0.5 \cdot I_{REQ}$ for $\sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = p_{\text{bias}}$. The presence of smaller undetected faults necessitates the second term.

- The second term is a product of the standard deviation of the estimation error (denoted by σ_{bias}) with a multiplier k_H. The derivation starts from the expansion for the error term ε_H, denoting the horizontal position error (Angus, 2006):

$$\begin{aligned}\varepsilon_H \coloneqq & \sqrt{\varepsilon_{(1)}^2 + \varepsilon_{(2)}^2} \\ = & \sqrt{\mu_{(1)}^2 + \mu_{(2)}^2} + \frac{\mu_{(1)}}{\sqrt{\mu_{(1)}^2 + \mu_{(2)}^2}} \left(\varepsilon_{(1)} - \mu_{(1)}\right) + \\ & \frac{\mu_{(2)}}{\sqrt{\mu_{(1)}^2 + \mu_{(2)}^2}} \left(\varepsilon_{(2)} - \mu_{(2)}\right) + \mathcal{O}\left(\sqrt{(\varepsilon_{(1)} - \mu_{(1)})^2 + (\varepsilon_{(2)} - \mu_{(2)})^2}\right),\end{aligned} \tag{2.74}$$

  with $\varepsilon_{(1)}, \varepsilon_{(2)}$ defined in Eq. 2.37 and $\mu_{(1)} = \mathbf{k}_{(1)}^T \mathbf{f},\ \mu_{(2)} = \mathbf{k}_{(2)}^T \mathbf{f}$.

  Ignoring the second and higher-order terms, ε_H is approximated by a Gaussian distribution $\varepsilon_H \sim \mathcal{N}(\sqrt{\mu_{(1)}^2 + \mu_{(2)}^2}, \sigma_{\text{bias}}^2)$. The variance σ_{bias}^2 is determined by

$$\begin{aligned}\sigma_{\text{bias}}^2 &= \mathbf{h}_H^T (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{h}_H, \\ \text{with } \mathbf{h}_H &= \begin{bmatrix} \frac{\mu_{(1)}}{\sqrt{\mu_{(1)}^2+\mu_{(2)}^2}} & \frac{\mu_{(2)}}{\sqrt{\mu_{(1)}^2+\mu_{(2)}^2}} & 0 & 0 \end{bmatrix}^T.\end{aligned} \tag{2.75}$$

  From Eq. 2.69 and Eq. 2.71, $\mathbf{h}_H$ is computed using

$$\begin{aligned}\frac{\mu_{(1)}}{\sqrt{\mu_{(1)}^2 + \mu_{(2)}^2}} &= \frac{\mathbf{K}_{1 i_{max}}}{\sqrt{(\mathbf{K}_{1 i_{max}})^2 + (\mathbf{K}_{2 i_{max}})^2}}, \\ \frac{\mu_{(2)}}{\sqrt{\mu_{(1)}^2 + \mu_{(2)}^2}} &= \frac{\mathbf{K}_{2 i_{max}}}{\sqrt{(\mathbf{K}_{1 i_{max}})^2 + (\mathbf{K}_{2 i_{max}})^2}},\end{aligned} \tag{2.76}$$

  where i_{max} indicates the element whose associated measurement results in the maximum slope.

  This term captures the impact of the noise, showcased by the spread of the blue clouds in Fig. 2.9. Limiting the fraction of the clouds above the PL line to I_{REQ} leads to a k_H coefficient such that the resulting HPL can upper bound both the deterministic and noise terms. In Eq. 2.77, this fraction is denoted by P_{noise}:

$$\begin{aligned}P_{\text{noise}} \coloneqq & P\left(\varepsilon_H - E(\varepsilon_H) \leq \text{HPL} - \text{HSLOPE}_{\max} \cdot p_{\text{bias}} \;\middle|\; \sqrt{\mathbf{f}^T \mathbf{S} \mathbf{f}} = p_{\text{bias}}\right) \\ = & I_{REQ}.\end{aligned} \tag{2.77}$$

  The knowledge about the estimation error distribution (Eq. 2.74 and Eq. 2.75) is used for the computation of k_H:

$$k_H = Q^{-1}(P_{\text{noise}}). \tag{2.78}$$

Consequently, the HPL is in the form of

$$\text{HPL} = \text{HSLOPE}_{\max} \cdot p_{\text{bias}} + k_H \cdot \sigma_{\text{bias}}. \tag{2.79}$$

To summarize, three probability specifications are utilized in computing the HPL: $P(\text{FA})$, $P(|q_{RB}| < T_{RB} \mid \sqrt{\mathbf{f}^T \mathbf{S} \mathbf{f}} = p_{\text{bias}})$ and P_{noise}. $P(\text{FA})$ is addressed by the CR requirement and the other two specifications are limited by the IR requirement. This setting ensures that the HMI probability will never exceed the required IR for *any* single-measurement fault condition. The proofs of safety are as follows (notably, the HMI probability in this section is concerning HPL, different from the one defined in Eq. 2.61 concerning AL):

- When p_{bias} is reached exactly, the conditional HMI probability is

$$\begin{aligned} P(\text{HMI} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = p_{\text{bias}}) =& P(|q_{RB}| < T_{RB} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = p_{\text{bias}}) \\ & \cdot P(|\varepsilon_H - E(\varepsilon_H)| > k_H \cdot \sigma_{\text{bias}} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = p_{\text{bias}}) \\ =& I^2_{REQ}. \end{aligned} \tag{2.80}$$

- In the case of greater fault magnitude ($\sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} > p_{\text{bias}}$), $P(\text{HMI} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} > p_{\text{bias}}) \leq I_{REQ}$ is easily known from Eq. 2.73.
- In the case of lower fault magnitudes ($\sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} < p_{\text{bias}}$),

$$\begin{aligned} P(\text{HMI} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} < p_{\text{bias}}) =& P(|q_{RB}| < T_{RB} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} < p_{\text{bias}}) \\ & \cdot P(|\varepsilon_H - E(\varepsilon_H)| > k_H \cdot \sigma_{\text{bias}} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = p_{\text{bias}}) \\ \leq& I_{REQ}. \end{aligned} \tag{2.81}$$

  The inequality is obtained by considering

$$P(|q_{RB}| < T_{RB} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = p_{\text{bias}}) \leq P(|q_{RB}| < T_{RB} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} < p_{\text{bias}}) \leq 1 \tag{2.82}$$

  and

$$\begin{aligned} P(|\varepsilon_H - E(\varepsilon_H)| > k_H \cdot \sigma_{\text{bias}} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} < p_{\text{bias}}) \\ \leq P(|\varepsilon_H - E(\varepsilon_H)| > k_H \cdot \sigma_{\text{bias}} \mid \sqrt{\mathbf{f}^T\mathbf{S}\mathbf{f}} = p_{\text{bias}}). \end{aligned} \tag{2.83}$$

The conditional HMI probability under fault-free hypothesis (H_0) can be bounded by

$$P(\text{HMI} \mid H_0) \leq P(|\varepsilon_H| > \text{HPL} \mid H_0) \leq I_{REQ}. \tag{2.84}$$

Consequently, the overall HMI probability is ensured.

Similarly, the VPL is computed using a counterpart $\text{VSLOPE}_{\max}$ and k_V:

$$\text{VPL} = \text{VSLOPE}_{\max} \cdot p_{\text{bias}} + k_V \cdot \sqrt{\left[(\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\right]_{33}}. \tag{2.85}$$

In practice, the IR requirement is allocated to the horizontal and vertical components arbitrarily (denoted by $I_{REQ,H}$ and $I_{REQ,V}$) depending on dedicated operations. The multipliers k_H and k_V are determined accordingly.

It is worth mentioning that this approach differs from Eq. 2.67, where a worst-case fault that maximizes the HMI probability is evaluated. Using p_{bias} may lead to a loose bound for IR, which has been realized and discussed in various studies (Ober, 1998; Milner and Ochieng, 2011). However, searching for the worst-case fault is computationally expensive, preventing its usage in earlier GPS RAIM.

2.4.6 Advanced RAIM

Integrity risk evaluation

The baseline ARAIM algorithm adopts the MHSS method, where multiple SS test statistics are computed against respective fault modes. The IR is evaluated and bounded based on the following considerations over the fault-free hypotheses H_0 and each faulty hypothesis H_i, for $i = 1, ..., h$ (Joerger and Pervan, 2014):

Under H_0: An upper bound is established as follows:

$$P(|\varepsilon_0| > \ell \wedge |q_{SS,1}| < T_{SS,1} \wedge ... \wedge |q_{SS,h}| < T_{SS,h} \mid H_0) \leq P(|\varepsilon_0| > \ell \mid H_0). \tag{2.86}$$

Joerger and Pervan (2014) argues that this inequality establishes a tight bound, given the CR requirements allocated to FA (similar to Eq. 2.64). The evaluation is straightforward, as the distribution of ε_0 is fully defined in Eq. 2.51.

Under H_i: The following consecutive inequalities illustrate how to derive a bound for the probability in faulty cases:

$$\begin{aligned}
&P(|\varepsilon_0| > \ell \wedge |q_{SS,1}| < T_{SS,1} \wedge ... \wedge |q_{SS,h}| < T_{SS,h} \mid H_i) && (2.87)\\
\leq &P(|\varepsilon_0| > \ell \wedge |q_{SS,i}| < T_{SS,i} \mid H_i) && (2.88)\\
\leq &P(|\varepsilon_0| > \ell \mid H_i \wedge |q_{SS,i}| < T_{SS,i}) \cdot P(|q_{SS,i}| < T_{SS,i} \mid H_i) && (2.89)\\
\leq &P(|\varepsilon_0| > \ell \mid H_i \wedge |q_{SS,i}| < T_{SS,i}) && (2.90)\\
\leq &P(|\varepsilon_i| + T_{SS,i} \cdot \sigma_{\Delta,i} > \ell \mid H_i). && (2.91)
\end{aligned}$$

- The first inequality (Eq. 2.88) ignores knowledge of No Detection for all test statistics, except for the one specifically designed to detect H_i.
- The second inequality (Eq. 2.89) rewrites the original formulation in the conditional form.
- The third inequality (Eq. 2.90) is a relatively loose bound due to the fact $P(|q_{SS,i}| < T_{SS,i} \mid H_i) \leq 1$.
- The fourth inequality (Eq. 2.91) is based on the feature of SS detection:

  $$|\varepsilon_0| \leq |\varepsilon_i| + |\Delta_{SS,i}|, \text{ with } |\Delta_{SS,i}| = |\varepsilon_0 - \varepsilon_i|, \tag{2.92}$$

  and the facts that $|\Delta_i| = q_{SS,i}\sigma_{\Delta,i}$ and $q_{SS,i} \leq T_{SS,i}$ in the detection.
- The bound on the HMI probability is established as follows:

  $$P(\text{HMI}) \leq \sum_{i=0}^{h} P(|\varepsilon_i| + T_{SS,i} \cdot \sigma_{\Delta,i} > \ell \mid H_i) \cdot P(H_i) \leq I_{REQ}, \tag{2.93}$$

  with ε_i defined in Eq. 2.51 and $\sigma_{\Delta,i}$ in Eq. 2.54.

In practice, some fault modes that are very unlikely to happen may not be monitored for efficiency reasons. Hence, the sum of probabilities of 'not-monitored' fault modes is excluded from the exact assessment process and bounded by a constant value, denoted by P_{NM}. For example, the event of n_{NM} or more simultaneous satellite faults is replaced by an upper bound, cf WG-C ARAIM TSG (2016) A.VIII.2. The final equation is

$$\sum_{i=0}^{h} P(|\varepsilon_i| + T_{SS,i} \cdot \sigma_{\Delta,i} > \ell \mid H_i) \cdot P(H_i) \leq I_{REQ} - P_{NM}. \tag{2.94}$$

PL computation

Eq. 2.94 can also be used to derive the PL, which, consequently, provides an intuitive, spatial representation of the volume that is guaranteed to contain the true position with a probability higher than $1 - I_{REQ}$ (Joerger et al., 2014).

Replacing ℓ in Eq. 2.94 with PL, an equation is constructed for a specific hypothesis H_i:

$$\mathrm{PL}_i = Q^{-1}\left(\frac{I_{REQ,i}}{2}\right)\cdot\sigma_i + Q^{-1}\left(\frac{C_{REQ,i}}{2}\right)\cdot\sigma_{\Delta,i}. \tag{2.95}$$

Solving for each hypothesis H_i provides a respective PL_i. The final solution of PL can be determined by an optimization process involving all PL_i. The baseline ARAIM algorithm additionally captures the impact of nominal bias $\mathbf{b}_{nom}$ (cf. Sec. 2.4.2) by its ℓ_∞ norm:

$$\left|\mathbf{k}^T_{i,(m)}\right|\,\mathbf{b}_{nom} \quad \text{for } m = 1,2,3. \tag{2.96}$$

With a few more practical adaptions, the final formulations are:

- **Vertical Protection Level (VPL):** The IR is below the requirement allocated to the vertical component:

$$2\cdot Q\left(\frac{\mathrm{VPL} - \left|\mathbf{k}^T_{0,(3)}\right|\,\mathbf{b}_{nom}}{\sigma_{0,(3)}}\right) + \sum_{i=1}^{h} P(H_i)\cdot Q\left(\frac{\mathrm{VPL} - T_{SS,i,(3)}\cdot\sigma_{\Delta,i,(3)} - \left|\mathbf{k}^T_{i,(3)}\right|\,\mathbf{b}_{nom}}{\sigma_{i,(3)}}\right) = I_{REQ,V}\cdot\left(1-\frac{P_{NM}}{I_{REQ}}\right). \tag{2.97}$$

  As a reminder, the subscript "(3)" indicates the vertical state, and the same rule applies to HPL; the subscript "V" and "H" stand for Vertical and Horizontal PL, respectively.

- **Horizontal Protection Level (HPL):** The calculation begins with the two horizontal components, namely, $m = 1$ and $m = 2$:

$$2\cdot Q\left(\frac{\mathrm{PL}_{(m)} - \left|\mathbf{k}^T_{0,(m)}\right|\,\mathbf{b}_{nom}}{\sigma_{0,(m)}}\right) + \sum_{i=1}^{h} P(H_i)\cdot Q\left(\frac{\mathrm{PL}_{(m)} - T_{SS,i,(m)}\cdot\sigma_{\Delta,i,(m)} - \left|\mathbf{k}^T_{i,(m)}\right|\,\mathbf{b}_{nom}}{\sigma_{i,(m)}}\right) = \frac{1}{2}\cdot I_{REQ,H}\cdot\left(1-\frac{P_{NM}}{I_{REQ}}\right). \tag{2.98}$$

  To obtain the final HPL, the ℓ_2 norm of the two horizontal components is calculated:

$$\mathrm{HPL} = \sqrt{\mathrm{PL}^2_{(1)} + \mathrm{PL}^2_{(2)}}. \tag{2.99}$$

For more details and proofs of safety, readers of interest are referred to literature such as (Blanch et al., 2015a) and the baseline ARAIM algorithm description document (WG-C ARAIM TSG, 2019, 2022).

3

Distribution-free uncertainty modeling for GNSS pseudorange measurements

Uncertainty modeling and error bounding for observations are of vital importance for high-integrity GNSS applications before they are transformed into the state domain. In addition to classical stochastic approaches, the necessity of interval representation for uncertainty becomes evident due to the remaining systematics in GNSS measurements, as discussed in Sec. 2.2.

This chapter aims to tackle this challenge by developing practical approaches to realistic uncertainty modeling and error bounding for GNSS pseudorange measurements. These approaches are distribution-free, describing the uncertainty by intervals. The GNSS pseudorange measurement model will be presented first, discussing the major error sources that need to be accounted for. Subsequently, various error sources, including the residual tropospheric error, residual ionospheric error, and multipath error, will be investigated using interval-based methods. Real-world measurements will be used for validation and performance evaluation.

3.1 GNSS pseudorange measurement model

A GNSS receiver measures the apparent signal travel time from the navigation satellite to the user, which is equal to the geometric range divided by the speed of light. It generates a local signal code replica and estimates the misalignment with the received signal. The unambiguous apparent signal travel time is determined by combining the time shift (a measure of the apparent transit time modulo the code chip length), the number of complete code chips, complete code repeats, and additional information from the satellite's navigation data. Multiplying this time with the speed of light yields the apparent range, also known as *pseudorange*. Notably, these time or range measurements differ from the true signal travel time or true range. The difference is mainly caused by the receiver's and the satellite's clock offsets relative to the GNSS system time, among other errors and signal delays (Teunissen and Montenbruck, 2017). In this regard, the GNSS pseudorange measurement model is expressed as a non-linear observation equation in Eq. 3.1:

$$\begin{aligned} p_r^k(t) = &\sqrt{\begin{pmatrix} (x^k(t-\tau_r^k) - x_r(t))^2 \\ +(y^k(t-\tau_r^k) - y_r(t))^2 \\ +(z^k(t-\tau_r^k) - z_r(t))^2 \end{pmatrix}} + c\left(\delta t_r(t) - \delta t^k(t-\tau_r^j)\right) \\ &+ I_r^k(t) + T_r^k(t) + \epsilon_{MP,r}^k(t) + \epsilon_{NLOS,r}^k(t) + \epsilon_{noise}(t), \end{aligned} \tag{3.1}$$

where,

- $p_r^k(t)$: pseudorange measurement for satellite k and receiver r at epoch t;
- τ_r^k: apparent signal travel time;
- $x^k(t-\tau_r^k), y^k(t-\tau_r^k), z^k(t-\tau_r^k)$: coordinates of satellite k at time of transmission;
- x_r, y_r, z_r: coordinates of receiver r at time of reception;
- $c\,(\delta t_r(t) - \delta t^k(t-\tau_r^j))$: speed of light multiplied by the difference of the receiver's and satellite's clock offsets ("metric clock error");
- I_r^k: ionospheric delay;
- T_r^k: tropospheric delay;
- $\epsilon_{MP,r}^k(t)$: multipath error;
- $\epsilon_{NLOS,r}^k(t)$: Non-Line-of-Sight (NLOS) extra path delay (EPD);
- ϵ_{noise}: noise term and other remaining error.

Eq. 3.1 indicates major error sources that are critical to land navigation. Various error correction and mitigation techniques against single or multiple error sources have been developed so far, categorized by Schön and Kutterer (2006) into four groups:

- More or less sophisticated **correction models** reduce the amount of systematics depending on the specific requirements and applications of the users;
- **Algebraic combinations** of the observations include, e.g., the Ionosphere-Free Linear Combination, single or double differences, etc;
- **Additional model parameters** can be introduced and estimated for high-precision applications;
- Systematics can be integrated into the **stochastic model** in terms of correlations or individual observation weights.

This thesis investigates explicitly integrity monitoring for the standard positioning approach - *pseudorange-based Single Point Positioning (SPP)*, for which the first of the aforementioned approaches is commonly applied to correct, e.g., the tropospheric and ionospheric delays. The uncertainties in the observation domain that the correction models give rise to can be well-captured by intervals. The remainder of this chapter introduces the interval-based procedures for bounding *the residual tropospheric and ionospheric errors*, and discusses how the *pseudorange multipath error* can be handled by interval-based methods. Interested readers are recommended to read textbooks such as Teunissen and Montenbruck (2017), Kaplan and Hegarty (2017), Morton et al. (2021), etc,.

3.2 Residual tropospheric error

3.2.1 Introduction

The GNSS signal is refracted as it travels through the neutral atmosphere, inducing one of the primary sources for GNSS ranging error: the tropospheric delay. Its non-dispersive feature for the L-band GNSS signals prevents the possibility of forming linear combinations to eliminate it in navigation applications.

The tropospheric path delay T_r^k in units of *meter* [m] depends on the tropospheric refractivity n_T integrated along the signal propagation path l (from satellite k to receiver r):

$$T_r^k = 10^{-6} \int_k^r n_T(l)\,\mathrm{d}l. \tag{3.2}$$

In general, the tropospheric delay consists of two components: a hydrostatic (or dry) component depending solely on pressure, and a non-hydrostatic (wet) component depending on the water vapor profile (Saastamoinen, 1972; Davis et al., 1985). The hydrostatic component can be well-modeled empirically thanks to the relatively predictable behavior of the dry atmosphere constituents, which vary minimally over temporal and spatial scales of hours and kilometers, respectively (Hobiger and Jakowski, 2017). On the contrary, the water vapor profile is highly inhomogeneous, and the weather patterns can change rapidly, making the affected signal propagation effect challenging to model and compensate with empirical models.

To denote the tropospheric delay for GNSS observations at any arbitrary azimuths and elevations, the Zenith Path Delay (ZPD) and Slant Path Delay (SPD) are introduced. Readers may find the term Zenith Total Delay (ZTD) in the literature, which is equivalent to ZPD. The SPD is determined by multiplying the a priori zenith troposphere correction (ZPD) with an elevation-dependent scaling factor (the mapping function, denoted by $M_T(\theta)$):

$$SPD = ZPD \cdot M_T(\theta). \tag{3.3}$$

For the hydrostatic and non-hydrostatic components, the corresponding terms, Zenith Hydrostatic Delay (ZHD) and Zenith Wet Delay (ZWD), are defined, respectively. The ZPD is typically on the order of 2.4 m, with 90% of this due to the ZHD. Imperfect modeling of the Zenith Wet Delay (ZWD) can significantly contribute to the uncertainty in estimating or describing the ZPD.

In practice, various well-developed empirical troposphere models are widely applied to correct for the tropospheric delays. For example, the Saastamoinen model (Saastamoinen, 1972) is developed for optics and applied in a "blind" mode with a priori standard atmosphere, e.g., ISO2533 (ISO/IEC GUIDE 98-3:2008(E), 2008) or U.S. standard atmosphere (U.S. Standard Atmosphere, 1976), and the Global Pressure and Temperature model model GPT2 (Lagler et al., 2013); the Radio Technical Commission for Aeronautics (RTCA) recommends the UNB3 model by Collins and Langley (1997) in MOPS for the Wide Area Augmentation System (WAAS) users, which originates from the model by Askne and Nordius (1987) to estimate the wet delay (RTCA/DO-229D, 2006); these models are also adopted in GPT2w (Böhm et al., 2015) and GPT3 (Landskron and Böhm, 2018), fed with corresponding empirical meteorological models. The performance of these models must be evaluated, and the residual error must be bounded to ensure high-integrity navigation solutions.

Up until now, the residual tropospheric error has been treated stochastically in almost all bounding methods. For example, in aviation, MOPS indicates a maximum standard deviation of 0.12 m globally for its correction model in the zenith direction, which is currently in use in SBAS and ABAS (RTCA/DO-229D, 2006; WG-C ARAIM TSG, 2015, 2016). Researchers have applied the extreme value theory to model residual tropospheric errors (Rózsa, 2018; Rózsa et al., 2020), conducted statistical evaluations of correction models (Feng et al., 2020; Lai et al., 2023), developed models for terrestrial navigation (Narayanan, 2023) and explored time correlation modeling has also been by bounding the Autocorrelation Function (ACF) and Power Spectral Density (PSD) (Gallon et al., 2021b).

For interval solutions, Collins et al. (1999) analyzes radiosonde data for the UNB3 tropospheric delay model and indicates a range of $[-20, 20]$ cm that covers over 99.9925% profiles investigated, representing a conservative bound valid for the entire WAAS-served area; Schön

and Kutterer (2006) introduces the method of sensitivity analysis to assess the residual tropospheric error in terms of uncertainty intervals, which is adopted by Dbouk (2021). This approach has not been validated with real data. In response, this thesis aims to estimate realistic interval bounds for residual tropospheric errors using the sensitivity analysis method with real-world measurements. Different meteorological models are used as inputs to the Saastamoinen model in the comparative analysis. The remainder of the section is based on the initial results published in Su and Schön (2022b) and will summarize the subsequent extensions.

3.2.2 Methodology: Sensitivity analysis via interval arithmetic

Sensitivity analysis is a forward modeling approach to assess the uncertainty due to remaining systematic errors and can be applied to GNSS observations. To obtain interval solutions, the approach by Schön and Kutterer (2006) computes sensitivity coefficients through partial differentiation and subsequently determines an interval radius for the final uncertainty budget, cf Eq. 2.22. This method relies on the model being differentiable and assumes that uncertainties of influencing factors are fairly small, which may not always be valid.

Researchers investigate the method of sample-based sensitivity analysis, taking advantage of Monte Carlo simulation, in the stochastic context (Schwieger, 2007b). This approach is practical for a "black box" model and may be applied to interval-described uncertainties. However, it can be less attractive due to relatively higher computational demands in the presence of explicit mathematical expressions of the investigated model. To cope with these issues, this thesis proposes to implement the sensitivity analysis via natural interval arithmetic, as described below.

The model's uncertainty budget (interval value $[f]$) is expressed as the sum of all influence factors' contributions:

$$[f] \triangleq [\underline{f}, \overline{f}] = \sum_{i}^{n_d} [f_i], \tag{3.4}$$

with a lower bound ($\underline{f}$) and an upper bound ($\overline{f}$).

The sensitivity f_i of f with respect to the change of one specific influence factor d_i in its interval $[d_i] \triangleq [\underline{\Delta d_i}, \overline{\Delta d_i}] + d_i^* = [\underline{d_i}, \overline{d_i}]$ is determined by

$$[f_i] \triangleq f([d_i] \mid \mathbf{d}^*) - f(\mathbf{d}^*), \tag{3.5}$$

where,

- $\mathbf{d}^*$ denotes nominal values of the influence factors that are input to the functional model f. Hence, the evaluation result is with respect to the model's output $f(\mathbf{d}^*)$.
- $f([d_i] \mid \mathbf{d}^*)$ is the set image of $[d_i]$ under f, cf. Sec. 2.18, representing the variation of the models' output due to any potential changes of d_i within $[d_i]$.

There are two main advantages:

- Compared to the approach of real-valued operations on interval radii, the implementation via natural interval arithmetic allows for individual estimation of the lower bound and upper bound. Therefore, the resulting uncertainty intervals are not necessarily symmetric with respect to $f(\mathbf{d}^*)$.
- While this approach is particularly straightforward for monotonic functions, interval mathematics also allows for nonlinear, discontinuous, or nonmonotonic models. Should any readers be interested in further details of interval computations, textbooks such as Moore et al. (2009) are recommended.

3.2.3 Uncertainty of model influence factors

Characterization of influence factors This work investigates the impact of residual tropospheric error due to the application of the *Saastamoinen* model developed by Saastamoinen (1972), which shows that the tropospheric Zenith Path Delay (ZPD) can be calculated from surface meteorological parameters:

$$T_{ZPD} = \beta_1 \frac{(1 + \beta_2 \cos(2\Phi) + \beta_3 H)}{\cos(z)} \left[p + \beta_4 \frac{e}{T} + \beta_5 e - B \tan^2(z) \right] + \delta_r, \tag{3.6}$$

where T_{ZPD} is the range correction in *meter* (m), p is the pressure at the antenna site in *hectopascal* (hPa), T is the absolute temperature in *Kelvin* (K), e is the partial water vapor pressure in hPa, B and δ_r are correction terms, $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ are constants, and z is the apparent zenith distance, which is determined from the true zenith distance Z of the satellite by the formula $z = Z - \delta z$:

$$\delta_z = \frac{\alpha_1}{T} \tan(Z) \left(p + \frac{\alpha_2}{T} e \right) - \alpha_3 \tan(Z) \left(\tan^2(Z) + 1 \right) \frac{p}{1000}. \tag{3.7}$$

In Eq. 3.7 α_1, α_2, and α_3 are constant coefficients, and Z can be determined from the station height H in *meter* (m), station latitude Φ in *degree* (°), and satellite elevation θ in *radius*. All the primary variables and constants above-mentioned constitute the vector of influence factors $\mathbf{d}_{SAAST}$, cf. Eq. 3.4-3.5, for the Saastamoinen model:

$$\mathbf{d}_{SAAST} = [T, p, e, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, B, \delta_r, H, \Phi, \theta]^T .$$

Type B evaluation Any physical quantity should be given with a meaningful number of digits, depending on, e.g., the indicative resolution of measurement devices. The uncertainty due to *rounding errors*, described in Sec. 2.2, must be accounted for in the evaluation. The introduced method is used, allocating an uncertainty budget for rounding errors to all influence factors of the Saastamoinen model, denoted by $\mathbf{d}_{SAAST}$. This allocation corresponds to the radii of symmetric intervals, cf. Eq. 2.22. For example, Dbouk and Schön (2019) suggests $\mathrm{rad}([p]) = 0.5 \cdot 10^{-4}$ when the input value for surface pressure is given with 4 digits.

However, this method of evaluating uncertainty may sometimes prove insufficient and overly optimistic. The actual range of variation for certain factors could differ significantly in magnitude. For example, in the case of the Saastamoinen model, researchers have examined the uncertainty of a constant coefficient $\beta_1 = 2.277 \cdot 10^{-3}$. Davis et al. (1985) proposes an error bar of $0.5 \cdot 10^{-6}$, while Zhang et al. (2016) suggests a value of $2.2794 \cdot 10^{-3}$. Accordingly, the interval radius is determined as the difference between the suggested value and the original one plus the rounding error, $\mathrm{rad}([\beta_1]) = (2.4 + 0.5) \cdot 10^{-6}$, indicating the maximum range of variation of β_1.

Type A evaluation Additional care must be taken for the meteorological parameters T, p, and e, as they vary temporally and spatially, significantly influencing the model's output.

According to Feng et al. (2020), pressure measurements at ground level may not be as representative of the "true" mean surface pressure as those taken at higher levels, possibly due to turbulence. This introduces uncertainty to the Saastamoinen model. Therefore, it is proposed that long-term statistical analysis using on-site measurements be conducted to estimate their interval bounds.

Without loss of generality, the ISO2533 standard atmosphere is analyzed in this work, used as the a priori parameters to feed the Saastamoinen model. A sliding window is defined

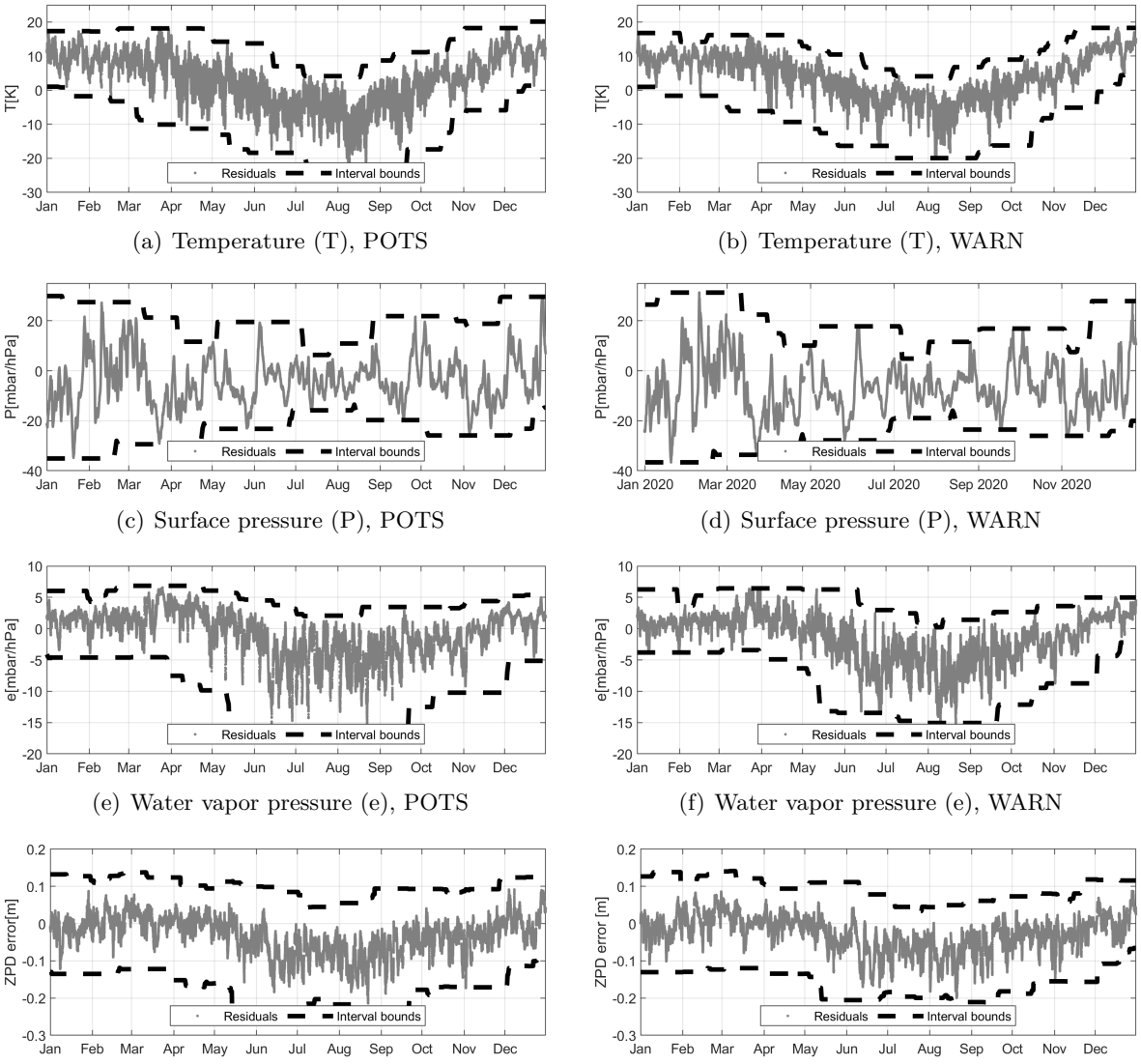
over two consecutive months in the time series to determine intervals that slide in daily steps. The residuals of parameters from the meteorological model with respect to the on-site measurements are computed within the window $t \in [t_1, t_2]$, using the equation:

$$\delta_\xi(t) = \xi(t) - \xi_0(t) \text{ for } t \in [t_1, t_2]. \tag{3.8}$$

Here, ξ denotes the outcome of the meteorological model (p, T or e) and ξ_0 is the corresponding on-site measurement.

Next, the upper and lower bounds of an interval are derived as minimum and maximum values. This allows for the daily interval bounds to be obtained, i.e., $[p]$, $[T]$, and $[e]$.

Fig. 3.1 shows the example results for the IGS stations Potsdam (POTS, a, c and e) and Warnemünde (WARN, b, d and f) during the year 2020: residuals of the ISO standard atmosphere (with respect to on-site measurements, shown as gray dots) and corresponding interval bounds (colored curves), which enclose all the residuals.



(a) Temperature (T), POTS

(b) Temperature (T), WARN

(c) Surface pressure (P), POTS

(d) Surface pressure (P), WARN

(e) Water vapor pressure (e), POTS

(f) Water vapor pressure (e), WARN

(g) ZPD error, POTS

(h) ZPD error, WARN

Figure 3.1: Example results of IGS station POTS (left) and WARN (right) in 2020: residuals (ISO-to-RNX) and bounds for meteorological parameters (temperature, pressure, partial water vapor) from long-term statistics. ZPD bounds are computed with these results and compared to residuals (SAAST-to-IGS)

Tab. 3.1 summarizes the influence factors and associated uncertainty evaluation in the current work.

Influence factors	Uncertainty evaluation
Temperature (T)	Type A: long-term observations
Pressure (p)	Type A: long-term observations
Water vapor pressure (e)	Type A: long-term observations
Model constant ($\beta_1 = 2.277 \cdot 10^{-3}$)	Type B: rad($[\beta_1]$) $= 2.9 \cdot 10^{-6}$
Model constants ($\alpha_1, \alpha_2, \alpha_3, \beta_2, \beta_3, \beta_4, \beta_5$)	Type B: rounding error by half of the last digit
Model correction terms (B, δ_r)	Type B: rounding error by half of the last digit
Station latitude (Φ) and height (H)	Type B: rounding error by half of the last digit
Satellite elevation (θ)	Type B: rounding error by half of the last digit

Table 3.1: Influence factors and associated uncertainty evaluation for the Saastamoinen model

3.2.4 Interval maps for residual tropospheric errors

Sensitivity analysis via interval arithmetic is applied to the Saastamoinen model after evaluating the uncertainty intervals for all influence factors $\mathbf{d}_{SAAST}$. The sensitivity of the model f_{SAAST} to each of the 16 elements of $\mathbf{d}_{SAAST}$, denoted by $[f_{SAAST,i}]$, is calculated using Eq. 3.5. Subsequently, the interval bounds for residual ZPD errors can be determined, denoted by $[f_{SAAST}]$, based on Eq. 3.4. At this stage, the impact of mapping functions for determining SPD is yet assessed.

To demonstrate results, ZPD residuals (δ_T), defined as the difference of computed ZPD (T_M) from the Saastamoinen model with respect to the reference estimates (T_R), can be used:

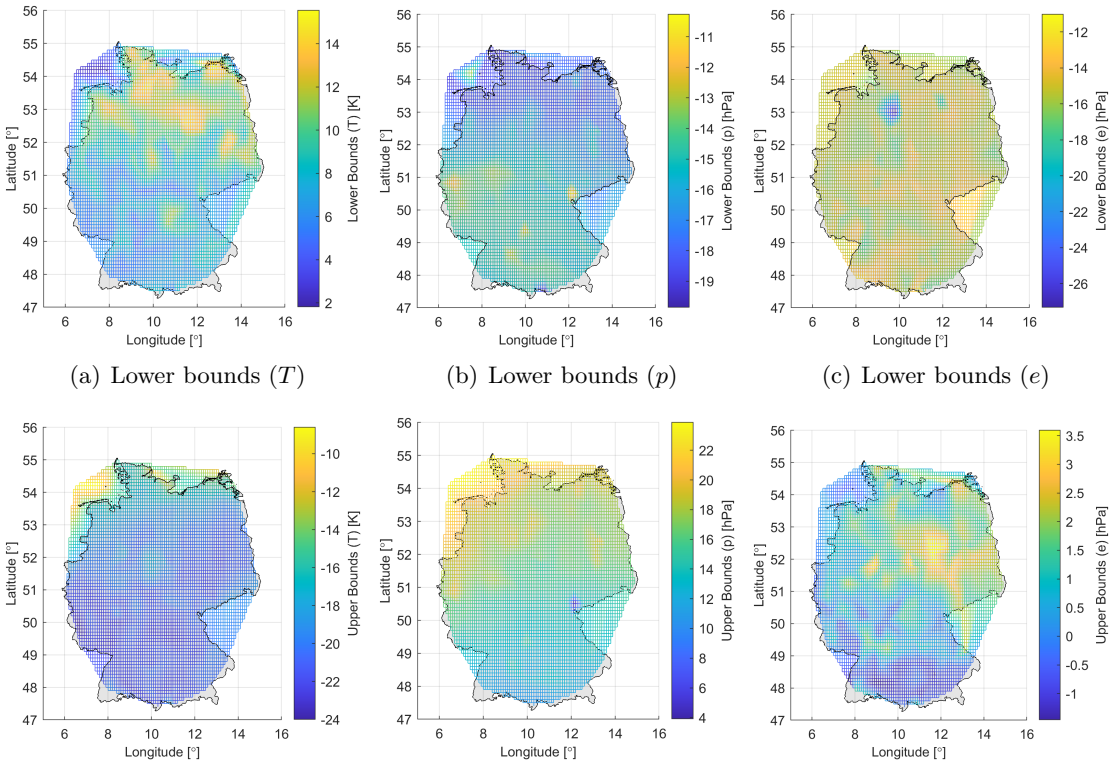
$$\delta_T = T_M - T_R. \tag{3.9}$$

By definition, the actual *residual ZPD error* is referenced to the truth. The IGS analysis centers generate ZPD products continuously for its global station network, with a reported accuracy of 4 mm (Johnston et al., 2017). This study considers the IGS ZPD products as sufficiently accurate for use as a reference. In Fig. 3.1 (g, h), the residuals are compared with the assessed interval bounds, showing a complete enclosure of gray dots as expected.

In addition, *Deutscher Wetterdienst (DWD)* operates a dense network of climate sensors over Germany (Kaspar et al., 2013), which enables the analysis for multiple stations and the estimation of error bounds across the country. Linear interpolation is used to create a grid network of 0.25° (latitude) ×0.25° (longitude) and obtained example results for three meteorological parameters on day 239 of 2020, as shown in Fig. 3.2.

Using these interval bounds as input to the proposed sensitivity analysis, the interval maps for residual ZPD errors are obtained cf. Fig. 3.3. The cross-sections of the interval maps in Fig. 3.3 along 9.7°E and 11.3°E meridians, which pass through Hannover and Potsdam, are presented in Fig. 3.4. From these figures, the focus is to identify potential dependencies of the uncertainty intervals for the residual ZPD error on meteorological parameters in terms of their geographical distributions. There are some noteworthy observations:

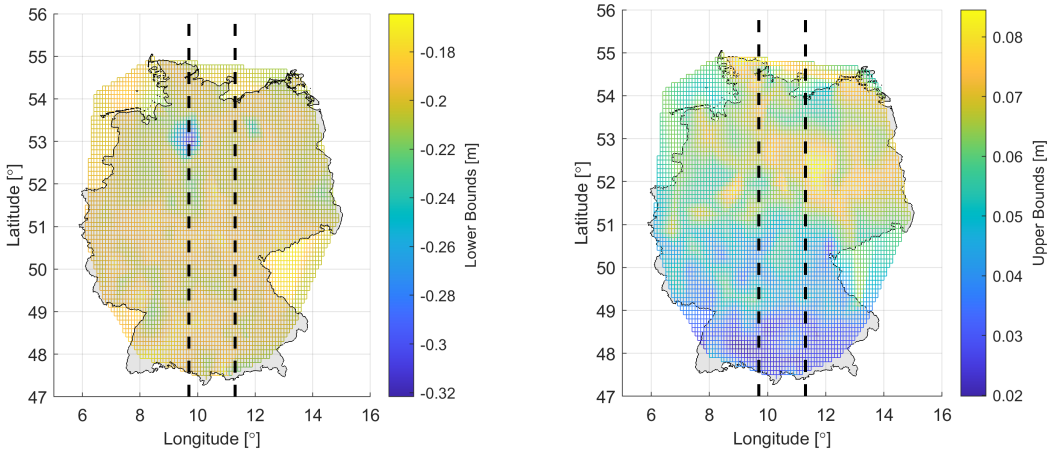
- The “wet” troposphere dominates the overall interval bound. Similar geographical distribution patterns are observed between water vapor pressure and residual ZPD error.



(a) Lower bounds (T) (b) Lower bounds (p) (c) Lower bounds (e)

(d) Upper bounds (T) (e) Upper bounds (p) (f) Upper bounds (e)

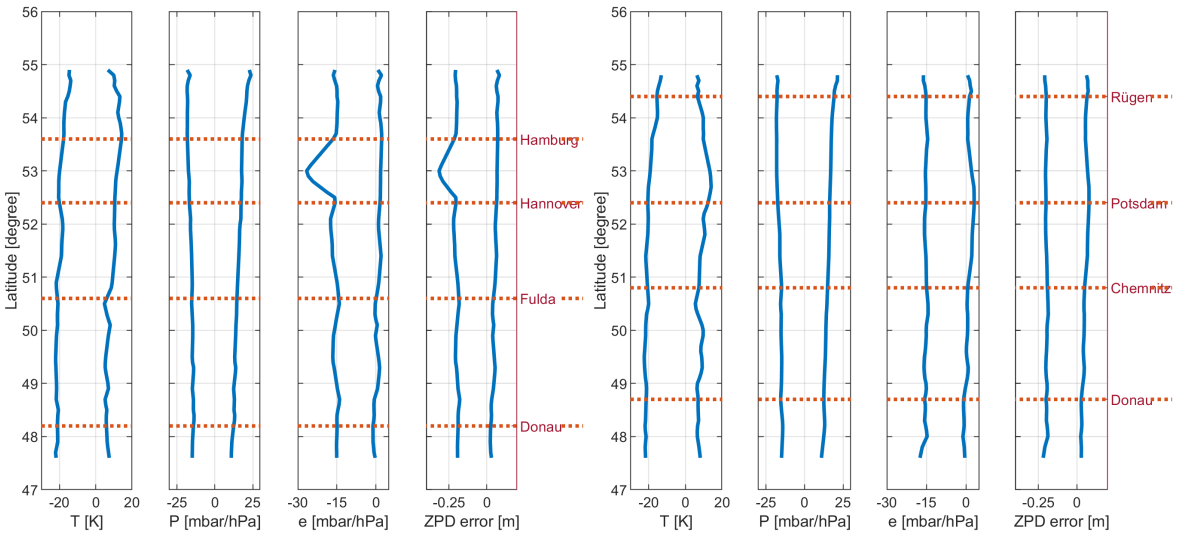
Figure 3.2: Geographical distribution of lower interval bounds (upper column) and upper interval (lower column) bounds (right) for meteorological parameters over Germany on DOY 239 in 2020. The parameters include the temperature (t), pressure (p), and partial water vapor pressure (e). The interval bounds are obtained from statistics for on-site measurements from 215 out of 345 DWD stations. Climate data source: *Deutscher Wetterdienst (DWD)* (Kaspar et al., 2013)



(a) Lower bounds (b) Upper bounds

Figure 3.3: Geographical distribution of lower interval bounds (left) and upper interval bounds (right) for the residual ZPD error over Germany on DOY 239 in 2020. The computation is based on the proposed sensitivity analysis for the Saastamoinen model using interval bounds from Fig. 3.2

- The impact of regional, small-scaled weather events, not captured by empirical tropospheric correction models, is accounted for by interval bounds. For instance, the region between Hamburg and Hannover indicates significantly wider intervals for water vapor pressure and residual ZPD error.



(a) Interval bounds along 9.7°E (b) Interval bounds along 11.3°E

Figure 3.4: Example results: Interval bounds for meteorological parameters (from ISO-to-DWD) and ZPD (Saastamoinen with ISO) along the 9.7°E (a) and 11.3°E (b) meridians (denoted in Fig. 3.3 as black dashed lines) on DOY 239 in 2020. Climate data source: *Deutscher Wetterdienst (DWD)* (Kaspar et al., 2013)

- The interval maps for residual ZPD errors exhibit good agreement with the station-wise experimental results. For example, there are consistent ZPD bounds for POTS station on day 239 of 2020, both when using on-site meteorological measurements (Fig. 3.1 (g), Oct. 26th) and using interpolated values from sensor network (Fig. 3.4 (b), 52.379°N). The differences observed in tests for the WARN station and for more days are at the level of mm to 10 mm.

3.2.5 Validation

This section is intended to validate the developed strategy through real-world scenarios, specifically the IGS station POTS, and the extension to further meteorological models:

- the UNB3 meteorological model (used in MOPS)
- the Global Pressure and Temperature model (GPT2w)

To quantify the overbounding performance, the bound-minus-residual (BMR) values, denoted by $\underline{\zeta}_T, \overline{\zeta}_T$, are defined as follows:

$$\underline{\zeta}_T = \underline{f} - \delta_T, \quad \overline{\zeta}_T = \overline{f} - \delta_T, \quad \text{with } \delta_T = T_M - T_R, \tag{3.10}$$

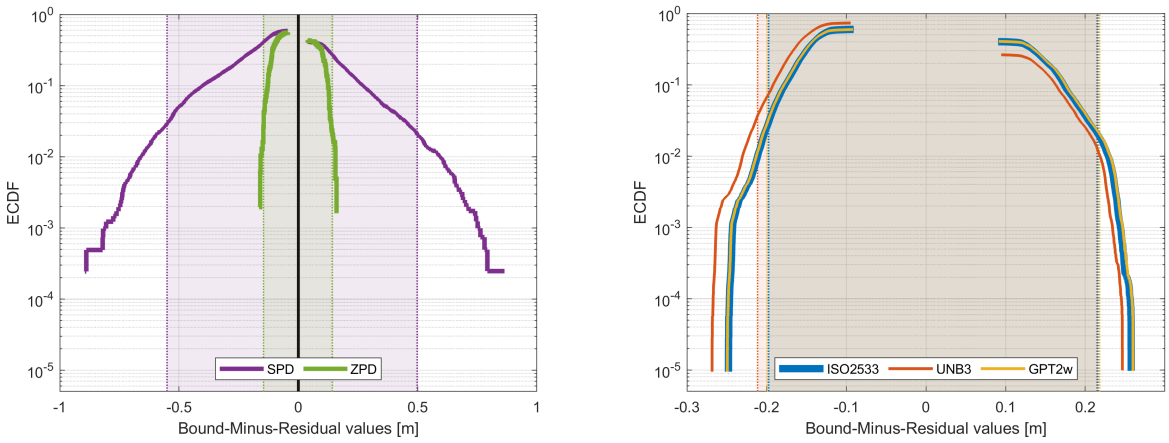
where $\underline{f}$ and $\overline{f}$ are the lower and upper bounds of the sensitivity analysis results, δ_T denotes the residuals of tropospheric delay, defined in Eq. 3.9 and extended to arbitrary directions. This indicator will be calculated in this section and used to evaluate the performance in the comparative analysis.

By definition, $\overline{\zeta}_T$ should be positive values and $\underline{\zeta}_T$ negative, while both are expected to be close to zero ideally. In this work, the Empirical Cumulative Density Function (ECDF) for a large amount of data is presented in a folded form (folded CDF): each residual corresponds to a single BMR value, either $\underline{\zeta}_T$ or $\underline{\zeta}_T$ depending on whichever has a smaller absolute value (i.e. which side the residual is biased toward); then, the CDFs of $\underline{\zeta}_T$ and $\underline{\zeta}_T$ are folded around zero and they sum up to one. Two key aspects are of interest:

- **Success of bounding**: Whether the bounds sufficiently enclose all residuals, i.e., non-negative $\overline{\zeta}_T$ and non-positive $\underline{\zeta}_T$.
- **Tightness of overbounding**: The width of the margin between the bounds and residuals in the case of successful bounding, i.e., the width of the folded CDFs.

Residual SPD errors Due to the difficulty of achieving ground truth for arbitrary tropospheric delays, this work takes estimates from Vienna Ray-Tracer as references (Hofmeister and Böhm, 2017). The ray tracing (RT) technique is capable of reconstructing the true signal path and calculating the atmospheric delay based on numerical weather models, providing high-quality estimates for ZPD and SPD. The online Ray Tracer allows for computing arbitrary SPD at IGS stations, requiring inputs of the Modified Julian Date (MJD), azimuth, and elevation (VMF Data Server, 2020).

Take the day 239 of 2020 as an example. The folded ECDFs of BMR for ZPD and SPD are provided in Fig. 3.5(a), with ZPD in green and SPD in blue. The left and right curves indicate $\underline{\zeta}_T$ and $\overline{\zeta}_T$, respectively, with the 95% intervals highlighted in light green and light purple, corresponding to the 5% quantile of $\underline{\zeta}_T$ and 95% of $\overline{\zeta}_T$. These intervals, along with the min/max range, are given in Tab. 3.2. It is worth noting:



(a) POTS station (b) WARN station

Figure 3.5: Folded ECDF of BMR values ($[\underline{\zeta}_T, \overline{\zeta}_T]$) for POTS station (left) and WARN station (right). For POTS station, the parameters of ISO standard atmosphere are input into the Saastamoinen model, computing zenith delays (ZPD) and slant delays (SPD) on day 239 of 2020. Residuals are obtained referencing the estimates from Vienna Ray-Tracer. ZPD at WARN station are computed for the same year with input parameters from ISO standard atmosphere, UNB3 model, and GPT2w model in the Saastamoinen model. Residuals are obtained referencing IGS ZTD products. For all curves in both sub-figures, 95% of the values are located within the respective shadow areas.

Table 3.2: 95% and min/max range intervals of folded ECDF for BMR values as in Fig. 3.5(a) (POTS station).

	95% interval [m]	**Min/max interval** [m]
ZPD	$[-0.146, 0.142]$	$[-0.165, 0.168]$
SPD	$[-0.551, 0.498]$	$[-0.895, 0.864]$

- The ECDF curves may not be perfectly symmetric around the y-axis due to (i) the imperfection of the reference and (ii) inadequate assessment of certain influence factors.

- While the proposed method assesses the uncertainty due to residual ZPD errors, the intervals for slant directions are computed by expanding ZPD intervals using mapping factors without directly assessing the impact of mapping functions on the uncertainty. Therefore, the ECDF curves for SPD BMR (blue) are significantly wider than those for ZPD. Nevertheless, the impact of mapping functions on the uncertainty may be observed from the change of asymmetry of ECDF curves from zenith to slant directions. For example, the min/max interval is "fatter" on the right side ($|\underline{\zeta}_T| \leq |\overline{\zeta}_T|$) for ZPD while "thinner" for SPD ($|\underline{\zeta}_T| > |\overline{\zeta}_T|$). This finding suggests the presence of systematic error due to mapping functions, which should be accounted for by an additional interval.

Further meteorological models In addition to the ISO standard atmosphere, the proposed strategy is validated for the UNB3 and GPT2w meteorological models. Unlike the other models, the UNB3 model provides parameters with respect to zero altitude above sea level instead of the Earth's surface. To minimize the impact of height discrepancies, the IGS station WARN, which has a geoid height of approximately 14 meters, is chosen. Fig. 3.6 (a/c/e) and (b/d/f) present the intervals for meteorological parameters from the UNB3 and GPT2w models. These intervals were obtained using the same procedures outlined in Sec. 3.2.3. The computed intervals for residual ZPD errors, determined through sensitivity analysis via interval arithmetic, are shown in Figures 3.6 (g/h).

Subsequently, BMR values are computed by referencing the Saastamoinen model outputs to IGS ZPD products. Their 95% and min / max intervals are summarized in Tab. 3.3.

Table 3.3: 95% and min / max range interval of folded ECDF for BMR values as in Fig. 3.5(b) (WARN station).

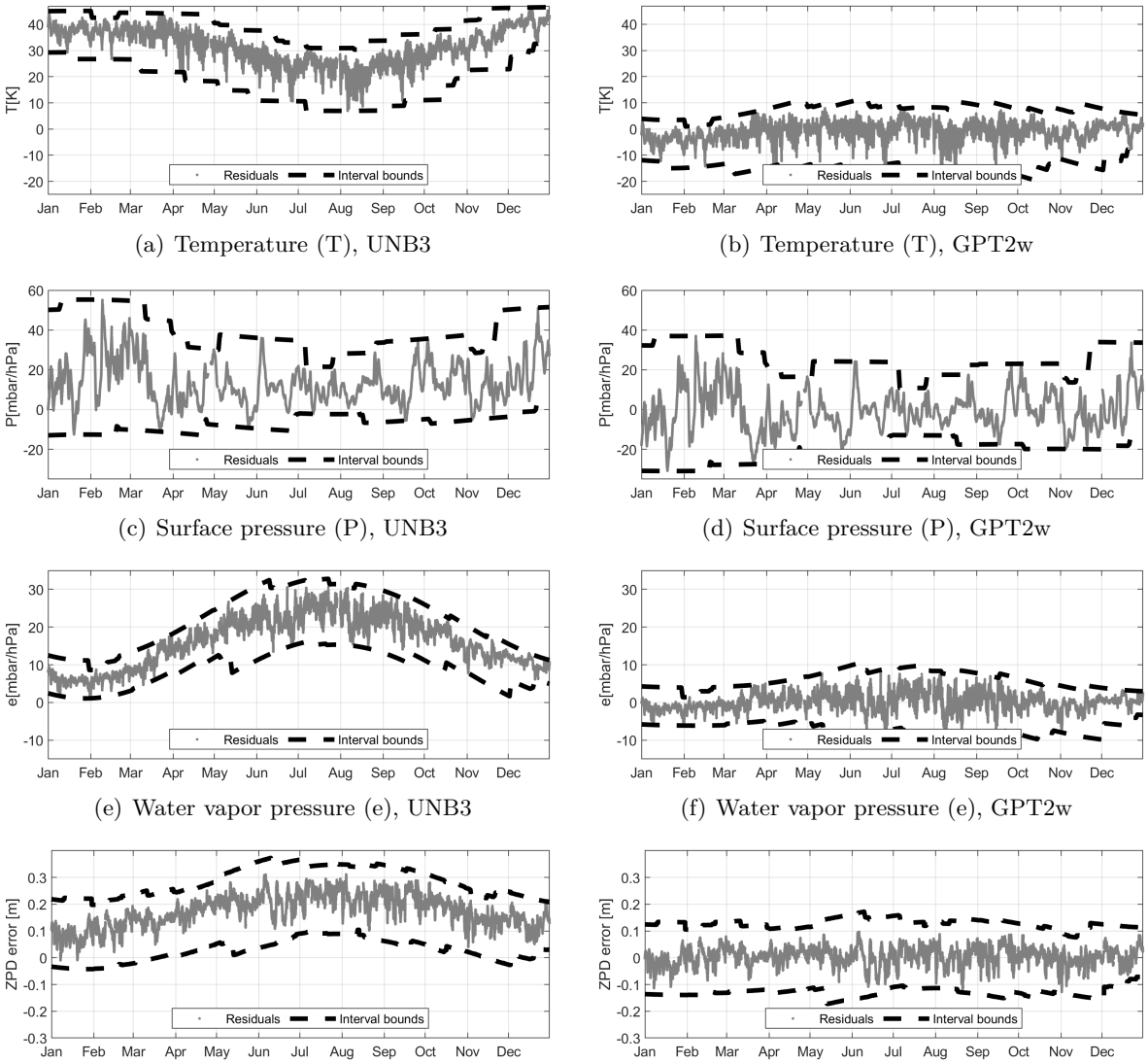
Meteorological model	**95% interval** [m]	**Min/max interval** [m]
ISO2533	$[-0.197, 0.216]$	$[-0.248, 0.258]$
UNB3	$[-0.211, 0.215]$	$[-0.269, 0.247]$
GPT2w	$[-0.199, 0.218]$	$[-0.249, 0.260]$

Key observations are as follows:

- The choice of meteorological model used in the Saastamoinen model has minimal impact on the uncertainty budget of residual tropospheric errors. This is because the inputs are representative values, and the model's construction process does not favor one model over another.
- High-quality meteorological models, such as GPT2w, contribute to more stable ZPD estimates, displaying better temporal consistency. The ZPD residuals appear symmetric around zero value, and the corresponding uncertainty intervals show less variation over months.

3.2.6 Discussion

The implementation of the Saastamoinen model demonstrates the feasibility of the proposed method for qualifying and bounding residual tropospheric errors. This is achieved through sensitivity analysis using interval arithmetic. The bounding performance was evaluated by calculating BMR values as overbounding indicators. Three empirical meteorological models were used as input to the Saastamoinen model, and tropospheric delay estimates from the IGS and Vienna Ray-Tracer were used as reference values. The results show that for the test



(a) Temperature (T), UNB3

(b) Temperature (T), GPT2w

(c) Surface pressure (P), UNB3

(d) Surface pressure (P), GPT2w

(e) Water vapor pressure (e), UNB3

(f) Water vapor pressure (e), GPT2w

(g) ZPD error, UNB3

(h) ZPD error, GPT2w

Figure 3.6: Example results of IGS station WARN in 2020: residuals and bounds for UNB3 (left) and GPT2w (right) meteorological parameters (temperature, pressure, partial water vapor) from long-term statistics. ZPD bounds are computed with these results and compared to residuals (SAAST-to-IGS)

data collected at the POTS and WARN stations, all ZPD and SPD residuals were successfully bounded, with 95% of the overbounding BMR values for ZPD around 0.2 m.

Careful assessment of the uncertainties in the model's influence factors is crucial. Meteorological parameters were evaluated through long-term statistics against on-site measurements, facilitating the modeling and bounding of seasonal and geographical dependencies. Taking advantage of a dense network of climate sensors, such as DWD, interval maps were generated to assess uncertainties in meteorological parameters and residual ZPD errors. Future improvements could include evaluating the impact of mapping functions to further reduce bounding conservatism.

This analysis is primarily intended for code measurements. For carrier-phase-based applications, more precise tropospheric correction models are used and, therefore, should be evaluated in future work. In highly complex models, the identification of numerous influence factors can pose a challenge, potentially limiting the applicability of the proposed method. Furthermore, the current implementation assumes independence between influencing factors. Any correla-

tions between these factors could lead to overestimated uncertainty, which represents another limitation of the method. Addressing the reduction of conservatism due to correlated factors should be a focus for future research.

3.3 Residual ionospheric error

3.3.1 Introduction

At approximately 1000 km above Earth's surface, GNSS signals encounter the ionosphere, a layer of electrically charged particles extending down to about 50 km altitude. Gas molecules in the ionosphere are broken up into free electrons and ions caused by the Sun's ultraviolet radiation. The induced ionospheric delay depends on the number of free electrons along the signal's path, described by the total electron content (TEC) (Hobiger and Jakowski, 2017):

$$TEC = \int_k^r n_e(l)\,\mathrm{d}l, \tag{3.11}$$

where TEC is defined as the number of electrons in a tube with a cross-section of $1\,\mathrm{m}^2$, and $n_e(l)$ is the electron density along the signal path integrated from the satellite to the receiver.

The electron density is not constant along the signal path due to spatial and temporal variations. This is because (i) different layers of the ionosphere are strongly affected by the ionization process in different ways, and (ii) the varying amount of solar radiation can cause diurnal, seasonal, as well as long-term (i.e., eleven years) variations in the electron density. In addition, unpredictable short-term effects due to irregular changes in solar activity and traveling ionospheric disturbances cause rapid changes in electron density (Hauschild, 2017).

The ionospheric delay is also frequency-dependent due to its dispersive nature. The phase and code of the same signal travel at different velocities, and hence, they must be treated separately. Approximating the refraction index to first order makes it possible to determine the ionospheric path delay, denoted by I_r^k, in the unit of length for pseudorange and carrier-phase measurements:

$$I_r^k = \begin{cases} \frac{40.3 \cdot TEC}{f^2}, & \text{for pseudorange measurement,} \\ -\frac{40.3 \cdot TEC}{f^2}, & \text{for carrier-phase measurement.} \end{cases} \tag{3.12}$$

In Eq. 3.12, f denotes the signal frequency. This equation reveals that the pseudorange delay has an opposite sign compared to the phase delay and decreases with increasing frequency, linearly depending on the TEC along the signal path to first order.

Dual- or multi-frequency receivers can significantly reduce or eliminate the ionospheric delay through ionosphere-free linear combinations. For single-frequency GNSS users, correction models such as the Klobuchar model for GPS (Klobuchar, 1987), NeQuick G model for Galileo (Nava et al., 2008), and BeiDou Global broadcast Ionospheric delay correction Model (BDGIM) model for BDS-3 are applied. These models rely on a set of model coefficients that are updated regularly and broadcast to users in the navigation message. However, these models only partially correct the ionospheric delay, for example, GPS Klobuchar model provides a correction capability of at least 50% root mean squared (RMS) error IS-GPS-200N (2022) and the Galileo NeQuick G model of 70% (European Commission, 2016), necessitating the evaluation of residual ionospheric errors for high-integrity applications. To this end, MOPS determines the variance of User Ionospheric Vertical Error (UIVE) for the GPS Klobuchar model and for the SBAS-based ionospheric correction model (RTCA/DO-229D, 2006); Schön

and Kutterer (2006) models the interval-represented uncertainty due to the residual ionospheric error through sensitivity analysis; however, it has not yet been validated with real data.

This section aims to investigate the feasibility of performing interval-based sensitivity analysis for bounding residual ionospheric error with real measurements, building on the initial results in Su and Schön (2021). The principle of interval-based sensitivity analysis is referred to Sec. 2.2 and Sec. 3.2.2. The remainder of this section will introduce the implementation and performance evaluation of a refined approach for residual ionospheric errors.

3.3.2 Characterization of influence factors for sensitivity analysis

The GPS-based *Klobuchar* model (Klobuchar, 1987) is the focus of this work. It approximates the diurnal variation of TEC as a positive half cosine-shaped curve with a constant bias (cf. Fig. 3.7). The phase of the cosine function is set at 14 h local time, and the vertical time delay at nighttime is set as a constant value of 5 ns. The amplitude and period of the cosine term are described as functions of geomagnetic latitude, represented by third-order polynomials, consisting of eight coefficients ($\alpha_i,\ \beta_i$ for $i = 1, 2, 3, 4$). These coefficients are updated at least once every six days by the GPS control segment and transmitted in the navigation message to users (IS-GPS-200N, 2022).

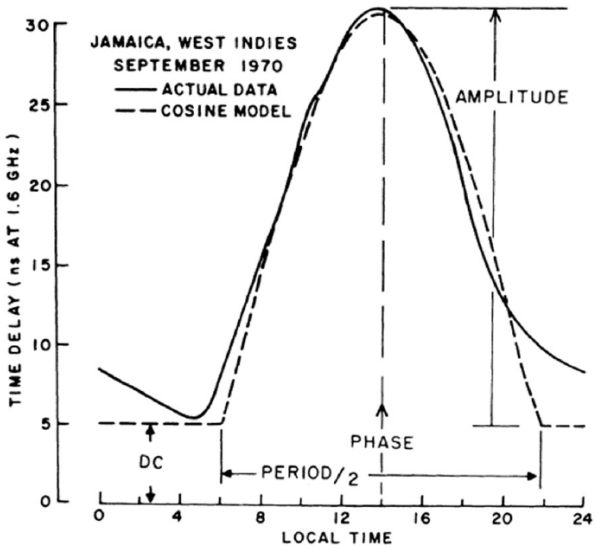


Figure 3.7: Cosine curve approximating the ionospheric effect in the Klobuchar model (Klobuchar, 1987)

The ionospheric correction algorithm is implemented in seven steps:

1. Calculate the Earth-centered angle ψ (semi-circles [1]):

$$\psi = \frac{0.0137}{\theta + 0.11} - 0.022, \text{ with } \theta \text{ the satellite elevation (semi-circles).} \tag{3.13}$$

   This is an approximation with an error of 0.2° to 0.4°, indicated by Klobuchar (1987). Accordingly, the influence factor concerning the approximation model f_ψ is associated with uncertainty represented by an interval of radius $df_\psi = 0.4°$.

2. Compute the subionospheric latitude Φ_I (semi-circles):

$$\Phi_I = \Phi + \psi \cos\phi, \text{ with } \phi \text{ the azimuth, } \Phi \text{ the user's latitude, both in semi-circle.} \tag{3.14}$$

[1] Conversion of units: 1 semicircle = 180° $(degree) = 0.5 \cdot \pi$ $(radius)$

If $\Phi_I > +0.416$, then $\Phi_I = +0.416$. If $\Phi_I < -0.416$, then $\Phi_I = -0.416$.

3. Compute the subionospheric longitude Λ_I:

$$\Lambda_I = \Lambda + \frac{\psi \sin\phi}{\cos\Phi_I}, \text{ with } \Lambda \text{ the user's longitude in semi-circle.} \tag{3.15}$$

4. Find the geomagnetic latitude Φ_m (semi-circles):

$$\Phi_m = \Phi_I + 0.064 \cos(\Lambda_I - 1.617). \tag{3.16}$$

This approximation is reported to represent the exact form to within 1° at all geomagnetic latitudes equatorward of 40°, and is within 2° up to 65° on either side of the geomagnetic equator. Hence, an influence factor concerning the approximation model f_{Φ_m} should be recognized, associated with an uncertainty interval of radius $df_{\Phi_m} = 1°$ or $df_{\Phi_m} = 2°$ depending on the geomagnetic location.

5. Find the local time t:

$$t = 4.32 \cdot 10^4 \Lambda_I + t_{GPS}, \text{ with } t_{GPS} \text{ the GPS time in second.} \tag{3.17}$$

If $t > 86400$, use $t = t - 86400$. If $t < 0$, add 86400.

6. Compute the slant factor $M_I(\theta)$ (unitless):

$$M_I(\theta) = 1.0 + 16.0 \cdot (0.53 - \theta)^3. \tag{3.18}$$

The uncertainty due to this approximated obliquity factor is up to two percent of the exact value. Hence, the associated uncertainty interval has a radius $dM_I = 0.02 \cdot M_I(\theta)$.

7. Compute the ionospheric time delay I (ns), referred to the L1 frequency:

$$I = M_I(\theta) \cdot \left[DC + \sum_{n=1}^{4} \alpha_n \Phi_m^{n-1} \cdot (1 - \frac{x^2}{2} + \frac{x^4}{24}) \right], \tag{3.19}$$

where

- $DC = 5 \cdot 10^{-9}$ represents the constant delay during nighttime;
- $x = 2\pi\,(t - t_0) \,/\, \left(\sum_{n=1}^{4} \beta_n \Phi_m^{n-1}\right)$ approximates the cosine term by third-order polynomials;
- $t_0 = 50400$ is the initial phase term, representing 14 h local time.

Wang et al. (2016) depicts that the nighttime delays may vary from 2.8 ns to 15 ns, depending on the latitude and solar activity. Accordingly, the DC term may be linked with an uncertainty interval of $[-2.2, 10]$ ns. The authors also investigate the model error induced by fixing the phase term to 14 h, which becomes significant near sunrise and sunset. The phase values mainly range from 13 h to 15 h as solar activities vary, based on which an interval of $[-1, 1]$ h may be assigned.

Analyzing the above model's construction process helps identify influence factors and associated uncertainties in the model:

$$\mathbf{d}_{KLOB} = [\theta, \phi, \Phi, \Lambda, t_{GPS}, t_0, DC, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2, \beta_3, \beta_4, f_\psi, f_{\Phi_m}, M_I] \tag{3.20}$$

Principally, the sensitivity of the model f_{KLOB} with respect to to each of the 18 elements of $\mathbf{d}_{KLOB}$, i.e., $[f_{Klob,i}]$, is determined using Eq. 3.5, given the uncertainty intervals of all

influence factors $\mathbf{d}_{KLOB}$. Subsequently, $[f_{KLOB}]$, the interval bounds for residual ionospheric errors can be computed based on Eq. 3.4.

Su and Schön (2021) conducted experimental analysis based on the above evaluation of $\mathbf{d}_{KLOB}$ and the method in Sec. 3.2.3, with test data from the IGS station Wettzell (WTZZ). The uncertainties for (partial) influence factors were assessed concerning their rounding errors in the absence of further information. It was found that the residual ionospheric errors were not adequately bounded by the obtained intervals, indicating an overly optimistic uncertainty evaluation. This means that intervals due to rounding errors can be insufficiently representative of these factors to reflect systematic uncertainty, and hence, the performance of the sensitivity analysis is degraded.

Researchers also investigated the eight broadcast coefficients, α_i and β_i for $i = 1, ..., 4$, whose uncertainty may be underestimated, because:

- The Klobuchar correction model can be characterized as a data-driven model, estimating the VTEC using ground-based GNSS data in an approximation model, hence, is never perfect in accuracy with respect to the actual values;
- The coefficients are selected from a set of pre-determined parameters based on the current date and solar activity, indicating limited representativity;
- The updating frequency is relatively low (i.e., daily), which may be insufficient to capture rapid, instantaneous events and anomalies.

To address these issues, two possible approaches may be used:

- Adding a *compensation interval* to the final interval budget as a supplementary component. Such an interval can be estimated based on the user's location and time using high-quality global ionosphere (TEC) maps (GIM). An introduction to computing compensation intervals will be given in Sec. 3.3.3.
- Inflating the intervals for influence factors with expert knowledge. Particularly, Sec. 3.3.4 proposes a method of evaluating the uncertainty intervals for the eight broadcast coefficients in the Klobuchar model despite lacking information from the system's operational side or initial determination of the coefficients. This is achieved by utilizing publicly available re-estimated products.

3.3.3 Estimation of compensation intervals

Estimating compensation intervals follows the principle of GUM Type A evaluation, i.e., taking advantage of long-term statistics for zenith ionospheric correction residuals. The reference values for zenith ionospheric delay are derived from IGS global ionosphere (TEC) maps products in the IONEX format. The residuals (δ_I) are subsequently computed as follows:

$$\delta_I = I_M - I_R, \tag{3.21}$$

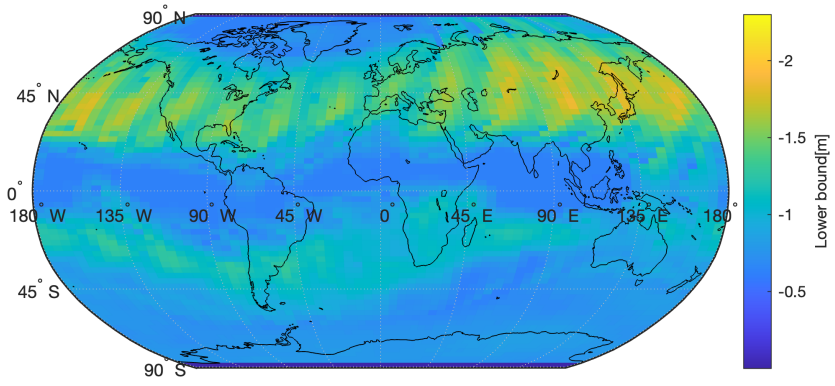
where I_M is the Klobuchar estimates, and I_R is the reference from IGS GIM.

To obtain daily intervals with respect to each ionospheric pierce points (IPP), a sliding window of two consecutive months (as an example) is defined along the time series of bound-minus-residual (BMR, denoted by $\underline{\zeta}_I$ and $\overline{\zeta}_I$):

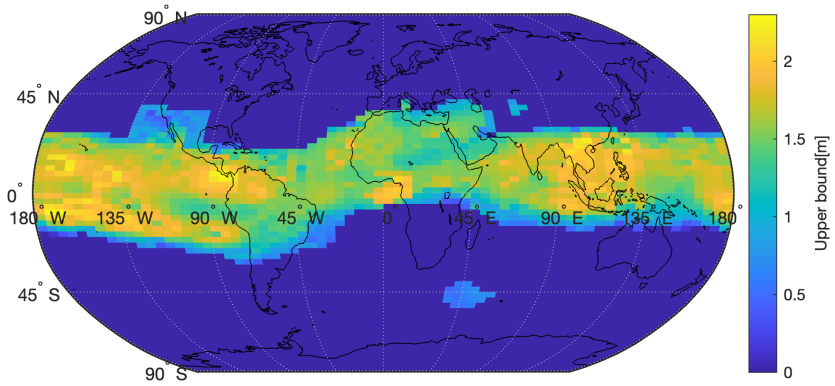
$$\underline{\zeta}_I = \underline{f} - \delta_I, \quad \overline{\zeta}_I = \overline{f} - \delta_I, \tag{3.22}$$

where $\underline{f}$ and $\overline{f}$ are the lower and upper bounds of the sensitivity analysis results. The objective is to determine compensation intervals $[\underline{C}_I, \overline{C}_I]$ such that $\delta_I \in [\underline{f}+\underline{C}_I, \overline{f}+\overline{C}_I]$. These intervals

are generated as the min/max enclosure of the sliding window. Fig. 3.8 shows an example of the lower bounds and upper bounds distribution for the ionospheric compensation intervals with respect to worldwide IPPs. These intervals are presented in grid nodes with a resolution of 2.5° (latitude) by 5° (longitude). Notably, these intervals are strongly geographically dependent, depicting significant upper bounds around the geomagnetic equator and the opposite situation for the lower bounds.



(a) Lower bound distribution of compensation intervals



(b) Upper bound distribution of compensation intervals

Figure 3.8: Global distribution of the compensation intervals computed for day 239 of the year 2020.

When analyzing a specific location, it is essential to consider all potential IPPs that could impact the received signals. The selection of IPPs is based on the orbital inclination and the satellite cutoff angle. For instance, the GPS, GLONASS, Galileo, and BDS MEO satellites orbit the Earth at inclinations of 55°, 64.8°, 56°, and 55°, respectively. Fig. 3.9 depicts the distribution of chosen IPPs (dashed red contour) for the Wettzell station and for GPS satellites when setting the elevation cutoff to 5°.

3.3.4 Uncertainty intervals for broadcast coefficients

To enhance the performance of the Klobuchar model, two approaches have been explored in the literature: (i) re-estimating the coefficients and constants in the model using higher quality GIM; (ii) introducing additional parameters for the Klobuchar model, such as replacing the constant *DC* term. Institutions like the Center for Orbit Determination in Europe (CODE) (Dach et al., 2009) and Chinese Academy of Sciences (CAS) (Wang et al., 2016) routinely re-estimate the eight broadcast coefficients, using different methodologies based on global TEC maps. These refinements offer stronger capabilities for correcting ionospheric delays compared to the broadcast coefficients in GPS navigation messages (Wang et al., 2019).

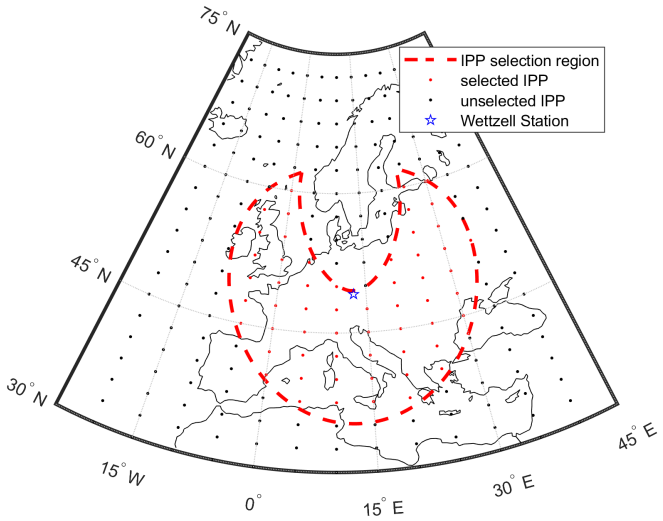


Figure 3.9: Example map of IPPs taken into the computation of compensation intervals for the Wettzell station and for GPS satellites. The grid resolution is 2.5° (latitude) × 5°(longitude).

This section is intended for a more realistic uncertainty budget. The variation range of the coefficients is critical, as any change that could improve the model's output indicates uncertainty in the coefficients. Consequently, comparing broadcast values to reliable re-estimated values yields intervals that indicate uncertainty. For instance, re-estimated values from CODE and their GIM products are used in the subsequent analysis (Dach et al., 2023). Figure 3.10 shows the residuals of the eight re-estimated coefficients relative to the GPS broadcast values. The interval bounds are computed as the enclosure of the residuals using a sliding window of two consecutive months. Additionally, the uncertainty due to rounding errors is taken into account to ensure consistent placement of intervals around the broadcast values.

A comparative analysis was conducted using data from 2020, examining residual ionospheric errors with respect to GIM for IPPs between $37.5°N$ to $60°N$ and $5°W$ to $40°E$. The success-bounding rate, the percentage of bounded residuals by sensitivity analysis results, increased from 52.8% to 72.5% when using the "improved" method instead of the "rounding" method, which solely accounts for rounding errors. The results, shown as histograms in Figure 3.11, indicate that while residuals around zero are mostly bounded in both cases, the "improved" approach better covers large residuals and partially improves the bounding of middle-sized residuals. However, the impact of remaining systematics is not fully captured by the current bounds. Possible reasons for this are:

- The re-estimated coefficients, while improving overall model performance, are still associated with uncertainties, and the resulting intervals are only partially representative;
- The GIM products used as reference have reported accuracies of 2-8 TECU in terms of RMS, which should be considered in the evaluation;
- Systematics due to other influence factors may persist and require further investigation.

Therefore, for residual ionospheric errors, a geographically and temporally dependent compensation interval is advised to be taken into consideration in applications.

3.3.5 Discussion

Implementing interval-based sensitivity analysis for the GPS Klobuchar model allows for determining interval bounds that represent the uncertainty due to the imperfect correction model.

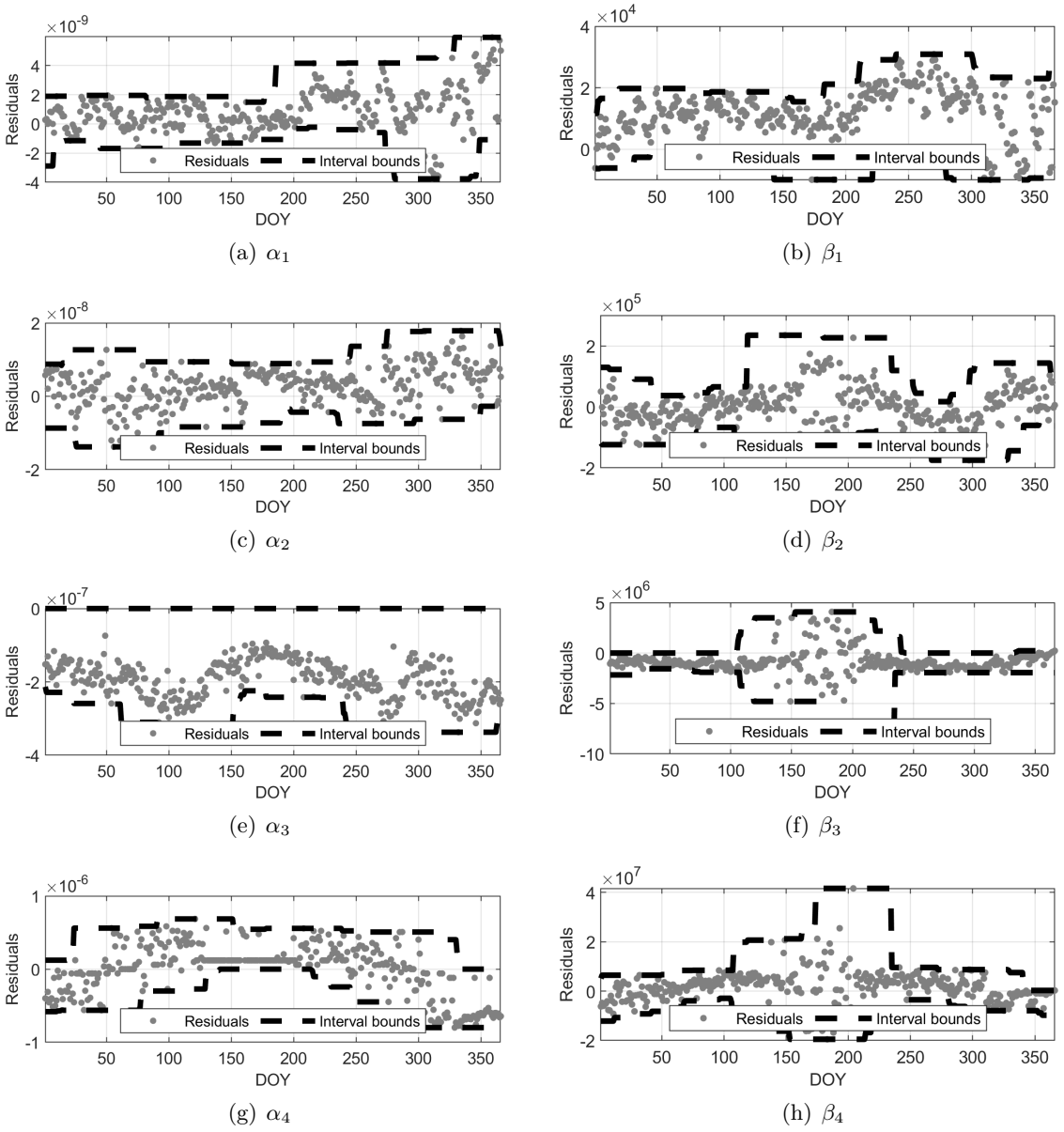
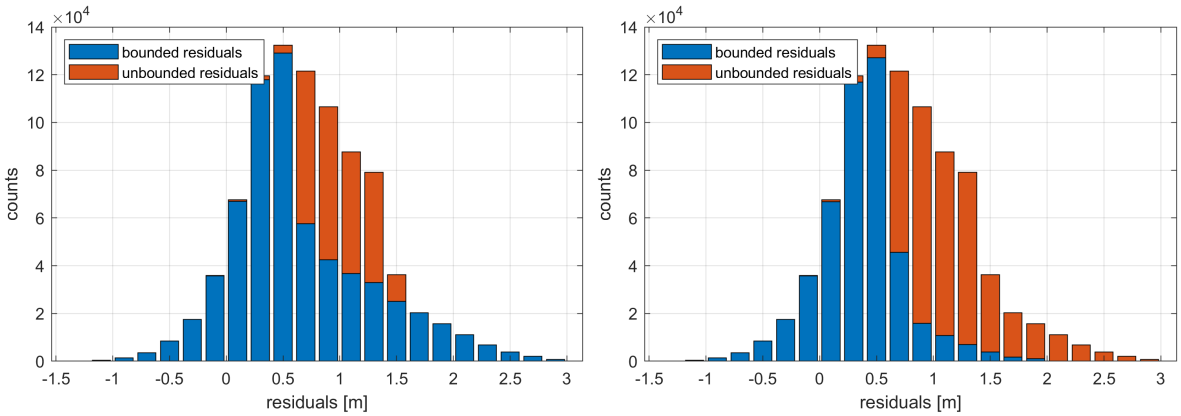


Figure 3.10: Intervals associated with the broadcast coefficients α_i and β_i for $i = 1, .., 4$: differences of the CODE re-estimated values from the broadcast values are observed (gray dots), from which the intervals are determined from their envelope. The rounding error is taken as the minimum bounds, assuring that the resulting intervals are placed around the broadcast values.

Two approaches have been evaluated to facilitate the bounding of residual ionospheric errors. By evaluating the uncertainty of GPS Klobuchar broadcast coefficients using re-estimated products, the bounding rate was significantly improved from 53.2% to 76.6% in an analysis for the WTZZ station. However, further investigation is essential to capture the remaining systematics to ensure high integrity. In the absence of complete information on partial influence factors, it is proposed to determine compensation intervals based on the user's location and time to guarantee full enclosure of residual ionospheric errors. Using IGS GIM products as a reference, compensation intervals can be derived considering selected IPPs. These intervals can be interpreted as supplementary components to the uncertainty budget.



(a) "improved" bounding results

(b) "rounding" bounding results

Figure 3.11: Histograms for residuals of vertical Klobuchar estimates with respect to the CODE GIM products. The data is collected from IPPs covering 37.5°N to 60°N and −5°W to 40°E for the year 2020. Two approaches are used to determine the uncertainty interval associated with the GPS Klobuchar broadcast coefficients: (a) derived from the CODE re-estimated coefficients ("improved"), (b) intervals representing the rounding error ("rounding"). Residuals successfully bounded by the sensitivity analysis results are indicated as black bars, and unbounded residuals as light gray bars in both histograms. The two-color bars do not overlap but sum up the overall histogram.

The evaluation of all aforementioned influence factors is summarized in Table 3.4, ensuring that residual ionospheric error can be safely bounded with intervals derived from sensitivity analysis.

Table 3.4: Influence factors and associated uncertainty interval evaluation for the GPS Klobuchar model.

Influence factors	Uncertainty evaluation
Satellite elevation (θ) and azimuth (ϕ)	Type B: rounding error by half of the last digit
Station latitude (Φ) and longitude (Λ)	Type B: rounding error by half of the last digit
GPS time (t_{GPS})	Type B: rounding error by half of the last digit
Phase (t_0) of the cosine term	Type A: historical data analysis
Constant delay during nighttime (DC)	Type A: historical data analysis
Broadcast coefficients $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2, \beta_3, \beta_4)$	Type B: rounding error and reference to reliable re-estimated products
Approx. of Earth-centered angle (ψ)	Type B: 0.4°
Approx. of geomagnetic latitude (Φ_m)	Type B: 1° or 2° depending on geomagnetic coordinates
Slant factor (mapping function) (M_I)	Type B: 2% of F
Compensation interval*$([\underline{C}_I, \overline{C}_I])$	Type A: statistical analysis referenced to GIM

* The compensation interval is not characterized as an influence factor but can be interpreted as a supplementary component to the uncertainty budget.

3.4 Multipath error in urban canyons

3.4.1 Introduction

Challenges that emerge in urban navigation include signal blockage, reflection, refraction, etc, due to sophisticated measurement environments. A prevalent issue is the *multipath effect*,

where GNSS signals arrive at the receiver via multiple paths due to reflection and diffraction. In this situation, the received signal is a combination of the direct Line-of-Sight (LOS) signal and Non-Line-of-Sight (NLOS) signals that are reflected and/or diffracted from nearby objects. The undesired NLOS component causes distortion of the received signal, leading to tracking errors in the receiver and resulting in pseudorange and carrier-phase observation errors, known as *multipath errors*.

Fig. 3.12 depicts a typical multipath scenario in a dense urban canyon environment, where the GNSS signal can arrive at the receiver via the direct LOS as well as through reflected and refracted paths.

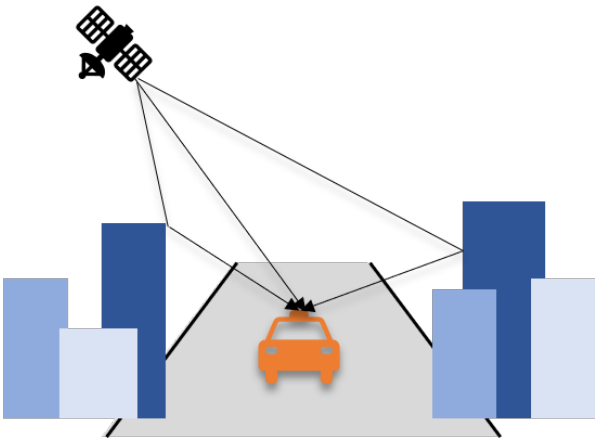


Figure 3.12: A typical measurement environment where the GNSS multipath effect occurs.

The multipath effect can have a significant impact on some applications. Pseudorange multipath errors can be as large as 100 m in severe conditions, while carrier-phase multipath errors can range from millimeters to centimeters (Braasch, 2017). To ensure accuracy, various techniques have been explored to model and mitigate multipath errors at the antenna, receiver, and navigation processor levels. Carefully designed antennas can significantly reduce multipath errors. Examples in the literature include antenna siting (McGraw et al., 2004), and choke ring technologies (Braasch, 2017), etc,. At the receiver level, techniques based on modified discriminator design have been developed to increase the resolution of the signal correlation process, thereby reducing the tracking errors. Additionally, vector tracking, which replaces the Delay Locked Loop (DLL) with an extended Kalman filter (EKF), has been demonstrated to track GNSS signals and calculate the user position simultaneously under various conditions (Pany et al., 2005; Hsu et al., 2015). 3D map aided GNSS is a typical technique deployed at the navigation processor level, utilizing city building models to improve the positioning accuracy through shadow matching, terrain height-aiding or NLOS detection (Adjrad and Groves, 2018; Ruwisch and Schön, 2022). Consistency checking can also be used to identify both NLOS and multipath-contaminated signals, although it is less effective in dense urban environments (Groves et al., 2013). In the context of cooperative positioning, vehicle-to-vehicle (V2V) ranging and vehicle-to-features (V2F) ranging can be leveraged to alleviate the impact of multipath effect and NLOS (Zhang et al., 2018; Soatti et al., 2018).

The main goal of this section is to upper bound the pseudorange error caused by multipath effects, as there is currently no standard correction model in place. This is especially important in urban areas where a large portion of the signals are affected by multipath because: (i) the fault detection and exclusion (FDE) function based on the self-consistency checking may not be effective; (ii) removing the erroneous signals can degrade satellite availability, threatening navigation continuity and availability. Therefore, the uncertainty budget, containing the multipath effect, should be adequately assessed for integrity purposes.

Stochastic approaches have been developed to address this problem. For example, the SBAS airborne multipath models leverage Gaussian overbounding concepts, providing a single-frequency model since 2000 (RTCA/DO-229D, 2006) and another for dual-frequency multi-constellation SBAS currently under development (Caizzone et al., 2022; Crespillo et al., 2024). There are also automotive multipath models based on similar concepts Khanafseh et al. (2018), as well as statistical models for urban canyon environments (Kong, 2011; Wang et al., 2018).

Alternatively, researchers have investigated the multipath error envelope (MEE) for GNSS signals – the upper and lower curves of the envelope represent the maximum delayed and advanced error on pseudorange measurements, respectively. They implicitly indicate the upper and lower bounds, and the actual values oscillate between these curves due to phase changes. The different signs depend on whether the multipath is in-phase or out-of-phase with the direct signal. Therefore, the interval representation is effective to upper bound the GNSS pseudorange multipath errors. Sec. 3.4.2 will introduce an interval bounding approach, followed by experimental validation and a discussion of its real-world application performance.

3.4.2 Methodology: Interval bounding based on multipath error envelope

Characterization of the multipath environment

In a typical multipath scenario like Fig. 3.12, all multipath signals travel longer distances than the direct signal, resulting in delayed arrival times. Typically, a multipath-affected signal is characterized by five parameters (Braasch, 2017):

- **Relative delay** is the *extra path delay* (EPD, denoted by D), induced by a multipath signal relative to the direct LOS signal. It can be determined using image theory and trigonometry.
- **Relative amplitude**, also known as the multipath-to-direct signal strength ratio (or M/D, denoted by ω), is dependent on the size, shape, and reflection coefficient of the multipath-inducing surface (i.e., reflector or diffractor). The reflection coefficient is a function of the incidence angle of the signal, typically approaching unity for near-parallel incidence. For example, Tab. 3.5 provides approximate values of the reflection coefficients and attenuation factors for common surface materials for GPS L1 frequency with normal incidence (elevation angle of 90°).

Table 3.5: Reflection coefficients and attenuation factors for common surfaces (normal incidence, GPS L1) (Braasch, 2017).

Surface material	**Reflection coefficient**	**Attenuation factor** (dB)
Soil (dry/moderate/wet)	0.268/0.566/0.691	-11.4/-4.94/-3.21
Glass	0.421	-7.51
Tinted glass	0.950	-0.446
Brick	0.345	-9.24
Concrete	0.404	-7.87

- **Relative phase** of the multipath relates to whether it interferes with the direct signal constructively or destructively. It is a function of the EPD and the reflection coefficient of the multipath-inducing surface and can change instantly upon reflection.
- **Relative phase rate**, also termed fading frequency, is related to the relative motion of the transmitter, receiver, and multipath-inducing obstacles. Notably, GNSS receiver

tracking loops can attenuate the multipath signal to some extent if its relative phase rate is higher than the tracking loop bandwidth.

- **Relative polarization**: GNSS signals are circularly polarized, but when reflected by a nonmetallic surface, they become elliptically polarized. This means that the reflected signal comprises both right- and left-handed polarized components. GNSS antennas are designed to receive the intended polarization and will reduce the strength of the opposite polarization. As a result, only the portion of the signal that becomes reverse-polarized due to reflection will be attenuated.

Multipath signal models

Van Nee (1993) modeled a received GNSS signal consisting of the LOS signal and n reflected signals:

$$s(t) = \sum_{i=0}^{n} a_i(t)\, m_c\left(t - \tau_i(t)\right) \cos\left(f_0\, t + \vartheta_i(t)\right) + \epsilon_c(t), \tag{3.23}$$

where

- i denotes the i-th component; the LOS signal component corresponds to $i = 0$;
- τ_i is the relative time delay of the i-th component;
- $a_i(t)$ is the amplitude of the i-th component. The relative amplitude of the i-th component with respective to the LOS component, i.e., M/D ratio, is $\omega_i(t) = a_i(t)/a_0(t)$;
- $m_c(t)$ represents the GNSS code modulation;
- f_0 is the nominal frequency of the LOS signal;
- $\vartheta_i(t)$ is the relative phase of the i-th component;
- $\epsilon_c(t)$ represents the noise.

The Delay Locked Loop (DLL) can track the code of most GNSS signals. However, multipath components can distort the correlation function, leading to tracking errors. This distortion causes pseudorange multipath errors, while the distorted phase of the received signals results in carrier-phase multipath errors. For a detailed description of the distortion process, readers of interest can refer to textbooks such as Braasch (2017).

Determination of multipath error interval bounds

The full extent of the envelope for a multipath signal was derived in Braasch (1997). The study considered the discriminator function and autocorrelation sidelobes for a binary phase shift keying (BPSK) signal such as GPS/GLONASS C/A-code and P-code. Accordingly, the upper bounds for multipath error can be derived based on the parameterization.

Eq. 3.24 shows the determination of MEE (denoted by ϵ_{MEE}) for a BPSK signal. The interval values $[\underline{\epsilon}_{MP}, \overline{\epsilon}_{MP}]$ can be computed as the upper and lower bounds of the envelope (ϵ_{MEE}), each consisting of four segments. The first segment is a slope, which is a function of M/D ratio and EPD, independent of correlator spacing and PRN chipping rate. In the extreme case of materials of the reflector, i.e., $\omega \to 1$, the lower slope will be infinitely close to vertical. The second segment depends purely on the M/D ratio and correlator spacing, regardless of the EPD. Therefore, the interval for the P-code within this segment will be significantly narrower than the one for the C/A-code. The third segment is additionally

affected by the discriminator, and the fourth segment is again independent of EPD.

$$\epsilon_{MEE} = \begin{cases} \dfrac{\omega_m D}{\omega_m + 1}, & \text{for } D \in \left[0, \dfrac{d_c}{2}(\omega_m + 1)\right]; \\ \dfrac{\omega_m d_c}{2}, & \text{for } D \in \left[\dfrac{d_c}{2}(\omega_m + 1), 1 + \dfrac{d_c}{2}(\omega_m - 1)\right]; \\ \dfrac{\omega_m(D - \frac{d_c}{2} - 1) \cdot (1 - \Gamma - \Gamma_{max} + \Gamma_{min}) - \omega_m d(\Gamma_{max} - \Gamma_{min})}{\omega_m(1 - \Gamma - \Gamma_{max} + \Gamma_{min}) - 2(1 - \Gamma)}, & \\ \qquad \text{for } D \in \left[1 + \dfrac{d_c}{2}(\omega_m - 1), 1 + \dfrac{d_c}{2}\left(1 + \dfrac{\omega_m(\Gamma_{max} - \Gamma_{min})}{1 - \Gamma}\right)\right]; \\ \dfrac{\omega_m d_c(\Gamma_{max} - \Gamma_{min})}{2(1 - \Gamma)}, & \text{for } D \in \left[1 + \dfrac{d_c}{2}\left(1 + \dfrac{\omega_m(\Gamma_{max} - \Gamma_{min})}{1 - \Gamma}\right), \infty\right]. \end{cases} \tag{3.24}$$

In this equation,

- ω_m has two possibilities: $\omega_m = \omega$ for the in-phase multipath signal (relative to the direct signal), and $\omega_m = -\omega$ for the out-of-phase case, with ω the M/D ratio.
- d_c represents the correlator spacing in the unit of PRN chips;
- Γ denotes the first sidelobe level of the autocorrelation function for PRN codes;
- Γ_{max} and Γ_{min} are the maximum/minimum autocorrelation sidelobe levels.

Newer signal structures have been developed based on binary offset carrier (BOC), multiplexed binary offset carrier (MBOC), composite binary offset carrier (CBOC) techniques, etc. The correlation functions of these signals have the same envelope as those of BPSK signals (with the same chipping rate) but differ via the sawtooth shape within the envelope (Braasch, 2017). As a result, the general BPSK multipath error envelope also encloses the envelopes of the aforementioned signals. For instance, Galileo E1 Open Service (OS) signals adopt the CBOC(6,1,1/11) modulation, the induced multipath error of which can be bounded by a simpler BOC(1,1) multipath error envelope. Without loss of generality, the analytical MEE model of a general BOC(n_1, n_2) signal is utilized. The resulting lower bound ($\underline{\epsilon}_{MP}$) or upper bound ($\overline{\epsilon}_{MP}$) is computed by (Harris and Lightsey, 2009)

$$\epsilon_{MEE} = \begin{cases} \dfrac{\omega_m D}{\omega_m + 1}, & \text{for } D \in \left[0, \dfrac{d_c}{2}(\omega_m + 1)\right]; \\ \omega_m \cdot (-1)^k \cdot d_c \cdot \dfrac{4\,\frac{n_1}{n_2} - 2\,k + 1}{2 - 8\,\frac{n_1}{n_2}}, & \text{for } D \in [D_{t_2,k}, D_{t_1,k+1}]; \\ \omega_m \cdot (-1)^{k+1} \dfrac{d_c\left(4\,\frac{n_1}{n_2} - 2\,k + 3\right) + \left(4(k-1) - 8\,\frac{n_1}{n_2}\right) \cdot \left(D + \frac{d_c}{2} - (k-1) \cdot \frac{T_{C/A}}{2\,n_1}\right)}{-8\,\frac{n_1}{n_2} + 2 + \omega_m \cdot (-1)^{k+1}\left(-8\,\frac{n_1}{n_2} + 4(k-1)\right)}, & \\ \qquad \text{for } D \in [D_{t_1,k}, D_{t_2,k}]; \\ \omega_m \cdot (-1)^{2\,n_1/n_2} \dfrac{-D + T_{BOC} + \frac{d_c}{2}}{-8\,\frac{n_1}{n_2} + 2 - \omega_m \cdot (-1)^{2\,n_1/n_2}}, & \text{for } D \in \left[D_{t_3}, T_{BOC} + \dfrac{d_c}{2}\right]; \\ 0, & \text{for } D \in [D_{t_1,k+1}, \infty], \end{cases} \tag{3.25}$$

where T_{BOC} is the service time associated with one chip of the BOC signal; n_1 and n_2 are normalized as multiples of the GPS C/A signal, with a chip rate at $1/T_{C/A} = 1.023\,\text{MHz}$.

Additionally,

$$\begin{cases} D_{t_2,k} = \omega_m \cdot (-1)^k \cdot d_c \cdot \dfrac{4\,\frac{n_1}{n_2} - 2\,k + 1}{2 - 8\,\frac{n_1}{n_2}} + \dfrac{d_c}{2} + (k-1) \cdot \dfrac{T_{BOC}}{2 \cdot n_1/n_2}, \\ D_{t_1,k} = \omega_m \cdot (-1)^{k+1} \cdot d_c \cdot \dfrac{4\,\frac{n_1}{n_2} - 2\,k + 3}{2 - 8\,\frac{n_1}{n_2}} + \dfrac{d_c}{2} + (k-1) \cdot \dfrac{T_{BOC}}{2 \cdot n_1/n_2}, \\ D_{t_3} = \omega_m \cdot (-1)^{2\,n_1/n_2} \cdot d_c \cdot \dfrac{n_2}{-8\,\frac{n_1}{n_2} + 2} - \dfrac{d_c}{2} + T_{BOC}, \\ T_{BOC} = \dfrac{T_{C/A}}{n_2}, \quad k = \left\lceil \dfrac{2D}{T_{BOC}} \cdot \dfrac{n_1}{n_2} \right\rceil, \end{cases} \tag{3.26}$$

where $\lceil \cdot \rceil$ is a ceiling operator. Subsequently, the interval values $[\underline{\epsilon}_{MP}, \overline{\epsilon}_{MP}]$ can be computed as the upper and lower bounds of the envelope (ϵ_{MEE}) (ϵ_{MEE}).

The subsequent experiments in this thesis apply this method to GPS L1 C/A (Eq. 3.24), GLONASS L1 C/A (Eq. 3.24), and Galileo E1 OS (Eq. 3.25-Eq. 3.26) signals. The basic parameters involved are given in Tab. 3.6.

Table 3.6: GNSS signal plan for GPS L1 C/A code, GLONASS L1 C/A code and Galileo E1 OS code.

	GPS L1 C/A	**GLONASS L1 C/A**	**Galileo E1 OS**
RINEX naming	GC1C	RC1C	EC1C
Central frequency	1575.42 MHz	(1598.0625 − 1605.375) ±0.511 MHz	1575.42 MHz
Spreading modulation	BPSK(1)	BPSK(0.5)	CBOC(6,1,1/11)
Primary PRN code length	1023	511	4092
Code frequency	1.023 MHz	0.511 MHz	1.023 MHz

3.4.3 Experimental validation

The multipath bounding method is applied to the dataset collected from a test drive in Hannover, Germany, in December 2021, detailed in Sec. 6.2.1.

Initially, example parameters in Eq. 3.24 and Eq. 3.25 are used to model the multipath error envelope. The parameters included:

- $\Gamma = 1/1023, \Gamma_{max} = 65/1023, \Gamma_{min} = -63/1023$ for GPS C/A-code according to Braasch (1997);
- $d_c = 1/30$ the early-minus-late (E-L) correlator spacing of the *Septentrio PolaRx5e* receiver for C/A-code (according to Hauschild et al. (2012) an older model *PolaRx2* of the product line, with "A Posteriori Multipath Estimator (APME)" inactivated)
- the M/D ratio ω determined by the incidence angle, antenna gains, and reflection coefficients that are associated with characteristics of the material. For example of the concrete, the conductivity is set $2 \cdot 10^{-5}$ S/m, the relative permittivity is around 3 F/m at GPS L1 frequency (Hannah, 2001).

  Mathematical models to determine the reflection coefficients, and subsequently the M/D ratio, can be found in Smyrnaios et al. (2013) and Icking et al. (2022).

Fig. 3.13 depicts an example of the multipath error envelope used in the experiment. Notably, the impact of autocorrelation sidelobes for the BOC signal is modeled similarly to Eq. 3.24 from a conservative perspective.

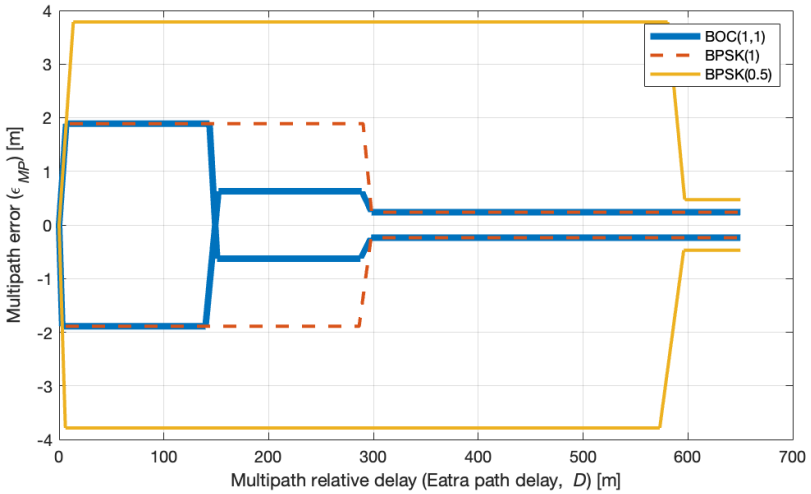


Figure 3.13: Example of multipath error envelopes for different signals used in the experiment: $\omega = 0.3877$, $d = 1/30$, $\Gamma = 1/1023, \Gamma_{max} = 65/1023, \Gamma_{min} = -63/1023$

Next, the ray tracer developed at IfE is employed to classify signals affected by the multipath effect (Icking et al., 2020, 2022). The ray tracer uses 3D city models to geometrically identify the signal obstructions, reflections, and refractions caused by buildings, walls, and other structures. For illustration, 256 epochs near Asternstraße are chosen, where the multipath effect is prominently observed (see Fig. 3.14(a)). The results of the signal classification from the ray tracer for all the tracked GPS signals are shown in Fig. 3.14(b). Six out of seven signals are classified as multipath affected at certain epochs and, therefore, should be accommodated by the proposed interval bounding method.

To validate the obtained multipath error bounds, the pseudorange multipath errors are computed using the ray tracer and used to form the Multipath Linear Combination (LC), cf., denoted by LC_{MP} in Eq. 3.27, for comparison:

$$LC_{MP} = p_1 - \varphi_1 + \frac{2\, f_2^2}{f_1^2 - f_2^2}\,(\varphi_2 - \varphi_1), \tag{3.27}$$

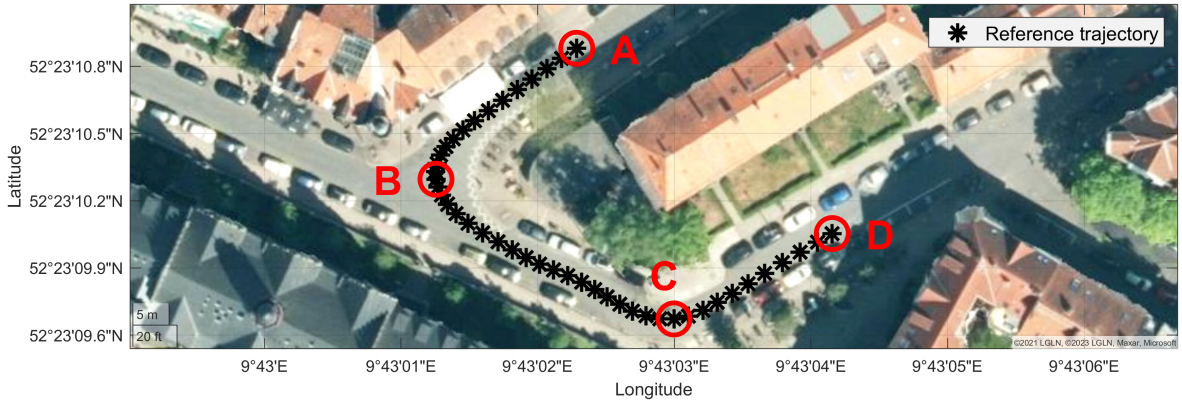
where p_*, φ_* are code and phase observations, with subscripts representing different frequencies, and f_* denoting the corresponding signal frequency.

The Multipath LC observation is potentially dominated by the pseudorange multipath error, apart from the noise terms and remaining higher-order ionospheric delays. The carrier-phase multipath error and carrier-phase noise are negligible due to their fairly low magnitude. Hence, the LC observations are expected to fall within the interval bounds.

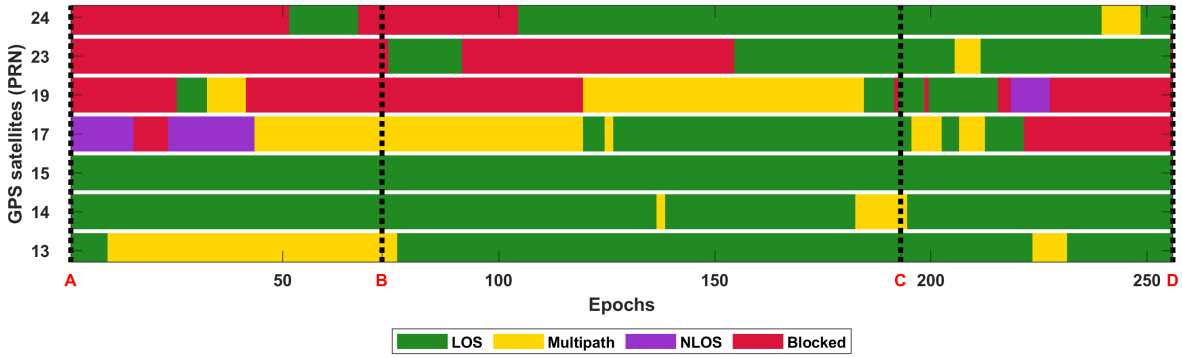
The remaining ambiguity term in the multipath LC should be removed first. One effective method to achieve this is to remove the median value in each continuous arc. Furthermore, cycle slips can frequently occur in challenging measurement environments, such as urban canyons, and need to be detected. When a cycle slip is found, the associated arc is split into two consecutive arcs and handled separately to mitigate the remaining ambiguities. This study uses a time-difference geometry-free LC method to detect cycle slips.

To minimize data gaps, three different LCs are formed, cf. Table 3.7 Line 1-3, involving GPS C1C and four phase observables: L1C, L2W, L2L, and L5Q (as RINEX 3 naming).

The experimental results are depicted in Fig. 3.15:



(a) Map of 'Asternstraße', where the experimental data was collected. The black dots indicate the reference trajectory and the driving direction from positions 'A' to 'D'.



(b) Signal classification by ray tracer developed at IfE (Icking et al., 2022). LOS signals are marked green, signals affected by multipath as yellow, NLOS as purple, and blocked signals as red. The ray tracer evaluates single reflections only, and hence, "blocked" signals may be tracked.

Figure 3.14: The measurement environment (a) and signal classification by ray tracer (b) for the data used in the multipath experiment.

- In Fig. 3.15(a), the computed pseudorange error from the ray tracer is plotted as colored dots for six GPS satellites (PRN 13, 14, 17, 19, 23, 24), with the corresponding interval bounds as dashed line segments. All the cases of interest present sufficient enclosure of errors as expected.
- In Fig. 3.15(b), three LCs are displayed with consistent color code and are absent at some epochs due to data gaps. Notably, the computed intervals successfully enclose the observed LC values for signals classified as multipath-affected.

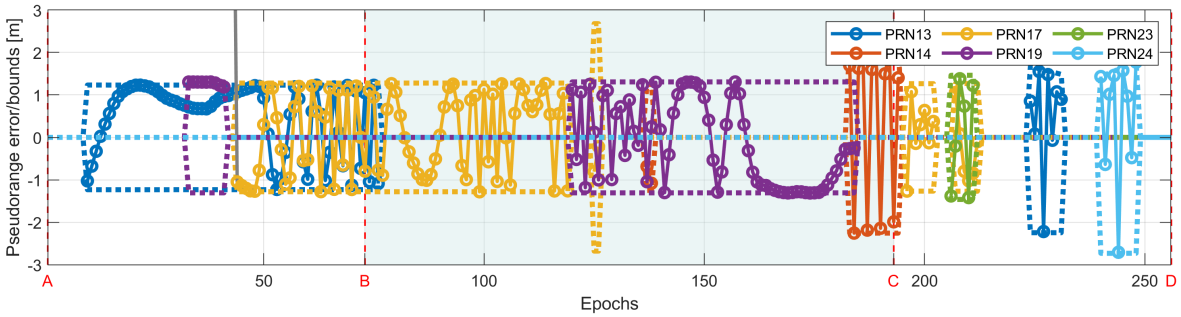
  The LC observations are typically smaller than the computed errors, possibly because: (i) suppression due to the relative polarization of reflected signals at the antenna level; (ii) the multipath component of the received signal may be better-attenuated thanks to the relatively high multipath phase rate at a moving receiver; (iii) the imperfect computation of multipath errors due to the uncertainty of the 3D city model used in the ray tracer.

For the analysis of the entire dataset of 9866 epochs, the BMR values (denoted by $\underline{\zeta}_{MP}$ and $\overline{\zeta}_{MP}$) can be calculated by subtracting multipath LC residuals from corresponding interval bounds (cf. Eq. 3.28) for GPS, GLONASS, and Galileo signals as Tab. 3.7. The results are presented as folded ECDF in Fig. 3.16.

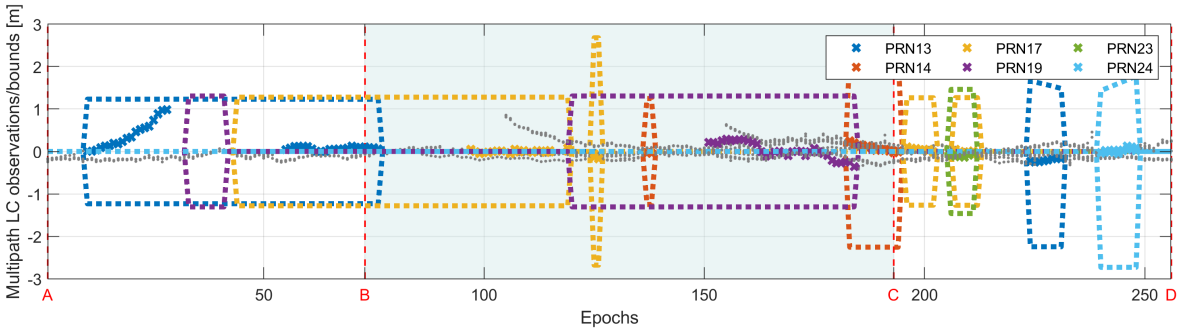
$$\underline{\zeta}_{MP} = \underline{\epsilon}_{MP} - LC_{MP}, \quad \overline{\zeta}_{MP} = \overline{\epsilon}_{MP} - LC_{MP} \tag{3.28}$$

LC	Code Obs	Phase Obs 1	Phase Obs 2	Bounded %	95% interval [m]
1	GC1C	GL1C	GL2W	93.258%	$[-1.308, 1.440]$
2	GC1C	GL1C	GL2L	94.213%	$[-1.430, 1.488]$
3	GC1C	GL1C	GL5Q	96.480%	$[-1.465, 1.537]$
4	RC1C	RL1C	RL2C	88.646%	$[-2.886, 2.581]$
5	RC1C	RL1C	RL3Q	95.805%	$[-3.007, 2.977]$
6	EC1C	EL1C	EL5Q	94.254%	$[-1.604, 1.492]$
7	EC1C	EL1C	EL6C	94.891%	$[-1.624, 1.518]$

Table 3.7: Seven multipath LCs formed in the experiment and corresponding bounding performance metrics: the percentages of bounded residuals and 95% range intervals of the BMR's folded ECDF.



(a) Computed pseudorange multipath errors using ray tracer and interval bounds.



(b) Multipath LC observations and interval bounds. Missing epochs are due to data gaps in phase observations.

Figure 3.15: Computed multipath errors (a, shown as "o") or multipath LC observations (b, shown as "x"), with corresponding interval bounds (dashed curves) from the experiment. Uniform colors are used for each satellite. Grey dots indicate (a) detours caused by NLOS and (b) observations not classified as multipath-affected by the ray tracer.

All ECDF curves crossing the y-axis indicate the presence of unbounded residuals, i.e., some LC observations are not bounded by the multipath interval bounds. Possible causes include: (i) receiver hardware noise, usually modeled using a Gaussian distribution and not captured by the multipath interval bounds; (ii) NLOS signals misclassified as multipath-affected due to the imperfect city model, leading to errors equal to EPD and exceeding the multipath interval bounds (also reported in Karimidoona and Schön (2023)); (iii) inaccurate estimates of EPD due to the limited quality of the city model, resulting in incorrect interval bounds from the MEE model; (iv) multiple reflections, which are not evaluated by the interval bounds. Additional key observations:

- Different LCs for the same code observable show similar performance in terms of tightness (similar patterns of BMR ECDF) and effectiveness (similar bounding rates). The GLONASS C/A code is exceptional, possibly due to insufficient observations.

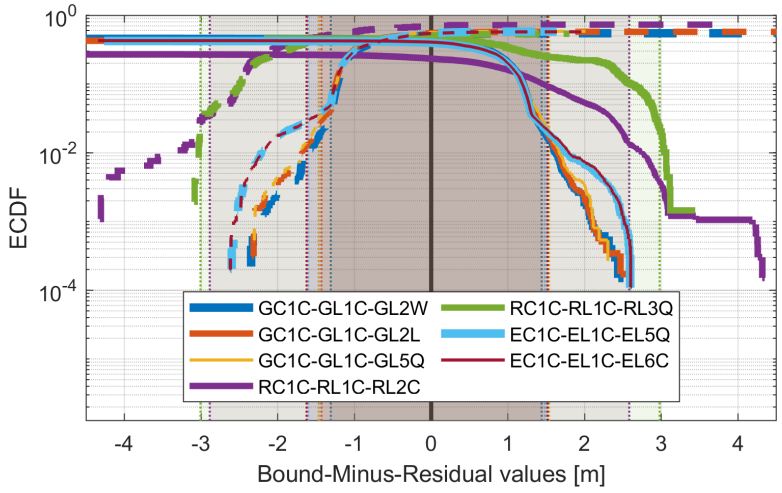


Figure 3.16: Folded ECDF of BMR values ($[\underline{\zeta}_{MP}, \overline{\zeta}_{MP}]$) computed from the investigated dataset. Seven multipath LCs are formed, for each of which $\underline{\zeta}_{MP}$ are plotted as dashed curves while $\overline{\zeta}_{MP}$ as solid curves. Corresponding 95% intervals are highlighted with consistent color code.

- The bounding performance for GPS is generally better than for GLONASS and Galileo. The GLONASS bounds are inherently wider due to signal modulation (see Fig. 3.13), whereas the BOC MEE model used in the computation is conservative for the actual Galileo CBOC signals, resulting in relatively loose bounds.

3.4.4 Discussion

Utilizing the multipath error envelope, interval bounds for the multipath errors can be determined with predefined parameters. The bounding performance depends on multiple factors: uncertainty of the building models, affecting the determination of multipath relative delay; uncertainty in parameters involved in multipath error envelope computation; Misclassification of signals, such as classifying multipath-affected signals into other categories, thus not evaluating the interval bounds. Based on the results presented, Sec. 6.2.3 will further investigate the impact of receiver hardware noise.

4

State estimation and error bounding

Estimation is a key concept across various fields, such as statistical inference, mathematical optimization, geodetic adjustment, and automatic control, all of which involve determining unknown parameters. In this chapter, the goal is to *estimate* the position and time as unknown states in the GNSS positioning and navigation problem, with a special emphasis on ensuring the reliability of state estimation for user safety (integrity purposes). In this regard, it is both reasonable and vital to address estimation errors through proper uncertainty modeling and its propagation in an acceptably conservative manner. This challenge motivates us to explore *error bounding* for state estimation, which constitutes the core objective of this chapter.

The content builds upon the foundational work presented in Su and Schön (2022c), which investigated uncertainty propagation and error bounding through a comparative analysis of the least-squares estimator, the interval extension of the least-squares estimator, and the set estimator based on constraint satisfaction. While the principles and methodologies introduced in that work form the basis of this chapter, the scope is significantly expanded to include additional point estimators and bounding strategies.

The chapter is organized as follows. The first section introduces the GNSS positioning problem. The second section revisits the classical least-squares estimator, which is widely used in GNSS applications, and discusses suitable error bounding approaches. The third section focuses on the set-based state estimation approaches, discussing the formulation of the set estimator based on constraint satisfaction and proposing two point estimators: the set-based central estimator and the set-constrained least-squares (SCLS) estimator. A comparative analysis of these estimators is presented, culminating in conclusive remarks about their applicability in GNSS tasks.

4.1 Introduction

The basic GNSS positioning problem aims to estimate the receiver coordinates ($\mathbf{p}_r = [x_r, y_r, z_r]^T$) and clock offset ($dt_r$) based on the non-linear observation equation in Eq. 3.1. In practice, providing an initial estimate of the receiver's position, denoted by $\mathbf{p}_{r,0} = [x_{r,0}, y_{r,0}, z_{r,0}]^T$, improves convergence. Taking advantage of Taylor's theorem, we can then linearize the geometrical term in the observation equation, neglecting the second and higher-order terms:

$$\rho_r^k \cong \rho_{r,0}^k + \mathbf{A}_p^k \Delta\mathbf{p}_r \tag{4.1}$$

where,

- $\Delta\mathbf{p}_r = \mathbf{p}_r - \mathbf{p}_{r,0}$ denotes the coordinate increment;

- $\rho_r^k = \sqrt{(x_r - x^k)^2 + (y_r - y^k)^2 + (z_r - z^k)^2}$ denotes the true range between the satellite k and receiver's position $\mathbf{p}_r$;
- $\rho_{r,0}^k = \sqrt{(x_{r,0} - x^k)^2 + (y_{r,0} - y^k)^2 + (z_{r,0} - z^k)^2}$ is the "computed" range between the satellite k and the initial estimate of the receiver r's position;
- $\mathbf{A}_p^k = \begin{bmatrix} \frac{\partial \rho_r^k}{\partial x_r} & \frac{\partial \rho_r^k}{\partial y_r} & \frac{\partial \rho_r^k}{\partial z_r} \end{bmatrix}$ reflects the *relative geometry* of the satellite k and the receiver r, forming a row vector of $\mathbf{A}_p$.

Notably, the uncertainty due to linearization, i.e., the error arising when neglecting second and higher-order terms, is considered negligible thanks to the considerably long distance between the GNSS satellites and the receiver on the Earth's surface.

The residual observations $y^k := p_r^k - p_{r,0}^k$, also termed Observed-Minus-Computed (OMC) values, are defined as the difference between the actual pseudorange observations (p_r^k, see Eq. 3.1) and the observations computed using the approximate parameter values and correction models ($p_{r,0}^k$). Consequently, the obtained measurement model is expressed as:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}, \tag{4.2}$$

where,

- $\mathbf{y}$: $[n \times 1]$, vector of measurements, i.e., the OMC values of pseudorange observations;
- $\mathbf{x}$: $[m \times 1]$, vector of states/parameters to estimate, including both the coordinate increment ($\Delta\mathbf{p}_r$) and receiver clock offsets (dt_r), e.g., $m = 4$ when using single satellite constellation;
- $\mathbf{e}$: $[n \times 1]$, measurement error vector, i.e., the remaining errors due to imperfect corrections or stochasticity, which should be modeled or bounded adequately;
- $\mathbf{A}$: $[n \times m]$, the design matrix. The row vectors in $\mathbf{A}_p$ from Eq. 4.1 are expanded to additionally capture the receiver clock offset, e.g., using GPS-only satellites:

$$\mathbf{A}^k = \begin{bmatrix} \frac{\partial \rho_r^k}{\partial x_r} & \frac{\partial \rho_r^k}{\partial y_r} & \frac{\partial \rho_r^k}{\partial z_r} & 1 \end{bmatrix} \tag{4.3}$$

$$= \begin{bmatrix} \frac{x_{r,0} - x^k}{\rho_{r,0}} & \frac{y_{r,0} - y^k}{\rho_{r,0}} & \frac{z_{r,0} - z^k}{\rho_{r,0}} & 1 \end{bmatrix} \tag{4.4}$$

This chapter addresses the core problem described above. Conventional solutions rely on stochastic methods, such as the family of the least-squares estimator. For instance, in civil aviation, RAIM and ARAIM make use of the weighted least-squares estimator. Critical considerations include the implementation of fault detection and exclusion functions, which involves the least-squares residuals, and the proper evaluation of estimation error bounds that need to be delivered to users (Brown, 1992; Joerger et al., 2014).

Additionally, several new estimators have been developed to fulfill different objectives. Robust estimators such as M-estimators can mitigate the impact of a small portion of outliers, hence enhancing robustness (Huber, 1981). The DIA estimator has been introduced to improve GNSS integrity through the least-squares estimation coupled with statistical testing (Teunissen, 1990; Teunissen et al., 2024). The Non-Least-Squares estimator deviates from the least-squares principle, trading off accuracy for improved integrity performance (Joerger et al., 2012). Blanch and Walter (2022) formulated an optimization problem concerning ARAIM integrity and continuity risks, which leads to a high-confidence region and a new estimator.

The interval-described uncertainty, in the literature, also known as set-membership or unknown but bounded uncertainty, offers an alternative approach to estimation problems, where

the optimality differs from the least-squares sense. The interval-described uncertainty is propagated to the state domain in a linear manner and, subsequently, represented or approximated by various types of sets. These sets are interpreted as the maximum possible variation in estimation errors. In this regard, the *deterministic* approach is more suitable for addressing *worst-case* scenarios, while traditional stochastic methods focus on an *average-case* sense (Casini, 2002). For example, in the field of navigation or localization, Drevelle and Bonnifait (2009) calculated high confidence regions for the vehicles' location as the union of numerous elementary boxes. Shi et al. (2017) investigated the wireless sensor network localization with bounded-error range measurements, and Dbouk and Schön (2018) explored different methods for computing bounding zones for unknown GNSS receiver positions.

Moreover, point estimators can be improved through interval-described uncertainty modeling. Milanese and Vicino (1991) provided a comprehensive overview of various classes of *deterministic* algorithms, in particular, the central algorithms and projection algorithms. Kacewicz et al. (1986) and Milanese (1995) connected the classical least-squares method with deterministic approaches, showcasing the optimality of the least-squares method under certain deterministic bounding conditions. Studies on filter techniques for a dynamic system with interval or set-described uncertainties are also investigated, for example, an extension of the classical Kalman filter to interval linear systems (Chen et al., 1997), zonotope-based methods (Combastel, 2003; Alamo et al., 2005), and a hybrid Kalman filter in Combastel (2016).

However, the potential benefits of set-based estimators for GNSS integrity have yet to be fully explored. To fill this gap, the remainder of this chapter covers different tasks toward four estimators, i.e., the classical least-squares estimator, set estimator based on constraint satisfaction, set-based central estimator, and set-constrained least-squares estimator. The organization of sections is summarized in Tab. 4.1.

Table 4.1: Organization of tasks in Chap. 4.

Estimator	**Task**		
	Esti. principle and properties	Error bounding approach	Performance evaluation and comparison
Least-squares estimator	Sec. 4.2.1	Sec. 4.2.2 - 4.2.4	Qualitative comparison: Sec. 4.4.1 Theoretical evaluation: Sec. 4.4.2 Statistical evaluation: Sec. 4.4.3
Set estimator based on constraint satisfaction	Sec. 4.3.1		
Set-based central estimator	Sec. 4.3.2		
Set-constrained least-squares estimator	Sec. 4.3.3		

4.2 Revisit: Error bounding for the least-squares estimator

4.2.1 Introduction

The basic idea of the *least-squares (LSQ)* method was described in detail by Gauss (1809):

> *The most probable value of the unknown quantities will be that in which the sum of the squares of the differences between the actually observed and the computed values multiplied by numbers that measure the degree of precision is a minimum.*

For the model in Eq. 4.2, the followings are presumed:

- The Variance-Covariance Matrix (VCM) of measurements is a positive definite matrix:

$$D(\mathbf{y}) = D(\mathbf{e}) = \mathbf{\Sigma} = \mathbf{P}^{-1}$$

.

- The expectation of measurement vector is a linear function of the unknown parameters:

$$E(\mathbf{y}) = \mathbf{A}\,\mathbf{x}$$

.

Let $\hat{\mathbf{y}}$ denote an estimator of $E(\mathbf{y})$, the least-squares method minimizes the quadratic form (Koch, 1999):

$$\Omega = (\mathbf{y} - \mathbf{A}\,\mathbf{x})^T \mathbf{\Sigma}^{-1} (\mathbf{y} - \mathbf{A}\,\mathbf{x}). \tag{4.5}$$

Subsequently, the least-squares estimator reads:

$$\hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P}\,\mathbf{y}. \tag{4.6}$$

Given sample values $\mathbf{y}$, the least-squares estimate of the parameters is obtained, along with

- The least-squares estimate of the measurements: $\hat{\mathbf{y}} = \mathbf{A}\,\hat{\mathbf{x}} = \mathbf{A}(\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P}\,\mathbf{y}$, and
- The least-squares residuals: $\mathbf{r} \coloneqq \mathbf{y} - \hat{\mathbf{y}} = (\mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P})\,\mathbf{y}$. Or, another notation $\hat{\mathbf{e}}$ is also found in literature, e.g., Teunissen (2003), interpreted as the least-squares estimate of the measurement errors.

For the sake of brevity, a $m \times n$ transformation matrix $\mathbf{K}$ is introduced to the remainder of the thesis: $\mathbf{K} = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P}$. Subsequently, the least-squares estimator in Eq. 4.6 can be expressed as $\hat{\mathbf{x}} = \mathbf{K}\,\mathbf{y}$.

Notably, the least-squares estimator applied in this chapter differs from that in Sec. 2.4.3, where it is used for fault detection by introducing an additional fault term to capture unmodeled or unbounded errors. Here, by contrast, the focus lies on state estimation under nominal conditions.

4.2.2 Stochastic error bounding

Fundamentally, the definition in Sec. 4.2.1 is established under two principles:

- The least-squares principle;
- The principle of best linear unbiased estimation.

The least-squares principle is a deterministic principle to the estimation problem without any statements concerning the characteristics of observations (Teunissen, 2003). It requires only the minimization of (weighted) sum of squared residuals. In contrast, the principle of best linear unbiased estimation relies on assumptions concerning the expectation and dispersion of the measurement error:

- $E(\mathbf{e}) = \mathbf{0}$;
- $D(\mathbf{e}) = \mathbf{\Sigma} = \mathbf{P}^{-1}$ is a known positive definite matrix.

Subsequently, the least-squares estimator gains some "optimal" properties in the sense of stochasticity:

- It is an unbiased estimator;

- It leads to the minimum variance of the estimated parameters.

The above properties are always true, independent of whichever distribution the observations follow. Concerning error bounding for the least-squares estimates, however, how to model the measurement error has to be specified.

For instance, in GNSS applications (as well as many other fields), the errors ($\mathbf{e}$) are usually assumed to follow zero-mean Gaussian distributions, or to be overbounded by inflated zero-mean Gaussian distributions in (advanced) RAIM. Based on this assumption, the least-squares method can provide the best linear unbiased estimation. The VCM of the estimates as a measure of their uncertainty is derived by quadratic variance-covariance propagation:

$$\begin{aligned}\boldsymbol{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} &= (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}\mathbf{A}(\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1} \\ &= (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}.\end{aligned} \tag{4.7}$$

The VCM contains all precision and correlation information, i.e., the variances (diagonal elements) and covariances between estimated parameters (off-diagonal elements). If we assume that the observation errors are normally distributed with VCM $\boldsymbol{\Sigma}$, then the estimates are normally distributed with VCM $\boldsymbol{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}$. The variances and covariances of the observations are obtained from the quadratic error propagation through the observation correction models, cf. Sec. 3.1.

With this, the concept of confidence region can be used to assess the uncertainty for the least-squares estimation. In particular, it is in the form of an ellipsoid in 3D and ellipse in 2D for the point position estimated from GNSS observations:

$$\mathcal{E}_{1-\alpha} = \{\mathbf{x} \in \mathbb{R}^m | (\mathbf{x}-\hat{\mathbf{x}})^{\mathbf{T}}\boldsymbol{\Sigma}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{\mathbf{-1}}(\mathbf{x}-\hat{\mathbf{x}}) \leq c_{1-\alpha,m}\}, \tag{4.8}$$

where $1-\alpha$ represents the confidence level and $c_{1-\alpha,m}$ is a constant, which can be determined by solving the following equation based on a chi-square distribution of m degree of freedom:

$$1-\alpha = \int_{-\infty}^{x} \chi^2(m,0)\,\mathrm{d}x. \tag{4.9}$$

4.2.3 Deterministic error bounding

The assumption of Gaussian distribution for observation errors can be easily violated, in particular, due to neglecting remaining systematics (Schön and Kutterer, 2005b) so that the confidence level no longer reflects the realistic uncertainty.

In order to assess the uncertainty due to remaining systematics, Kutterer (1994) proposed a straightforward method *the interval extension of the least-squares estimator*, i.e., replacing the real-valued measurement vector by an interval vector that is symmetric around zero and applying the computation rules from interval mathematics.

In the model of Eq. 4.2, a symmetric interval vector around zeros $[\mathbf{s}] = [-\boldsymbol{\Delta}_s, \boldsymbol{\Delta}_s]$ is constructed to represent the uncertainty due to remaining systematics in the observation domain. Subsequently, the uncertainty is transferred linearly to the state domain by the $m \times n$ matrix $\mathbf{K} = (\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{P}$:

$$\left((\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{P}\right)\,[\mathbf{s}] = \mathbf{K}\,[\mathbf{s}]. \tag{4.10}$$

In this way, the uncertainty due to remaining systematics in the state domain can be represented by Eq. 4.10 in the form of a set. Given measurement values $\mathbf{y}$, the position information

is related:

$$\begin{aligned}\mathcal{Z}_D =&\{\mathbf{x}\in\mathbb{R}^m \,|\, \mathbf{x}\in\mathbf{K}(\mathbf{y}-[\mathbf{s}])\}\\ =&\{\mathbf{x}\in\mathbb{R}^m \,|\, \mathbf{x}\in\mathbf{K}\,\mathbf{y}-\mathbf{K}\cdot[-\boldsymbol{\Delta}_s,\boldsymbol{\Delta}_s]\}\\ =&\left\{\mathbf{x}\in\mathbb{R}^m \,|\, \mathbf{x}=\hat{\mathbf{x}}+\sum_i^n w_i\cdot\mathrm{col}_i(\mathbf{K})\cdot\Delta_{s,i},\ \text{with}\ -1\le w_i\le 1\right\},\end{aligned} \tag{4.11}$$

where $\mathrm{col}_i(\mathbf{K})$ represent the i-th column vector of $\mathbf{K}$.

By definition, the set in Eq. 4.11 is a zonotope $\mathcal{Z}(\hat{\mathbf{x}},\mathbf{K}\cdot\mathrm{diag}(\boldsymbol{\Delta}_s))$ with its center being the least-squares solution $\hat{\mathbf{x}}$ and generators being $\mathbf{K}\cdot\mathrm{diag}(\boldsymbol{\Delta}_s)$. It is a suitable measure of point uncertainty due to remaining systematics that is linearly propagated (by $\mathbf{K}$) from the observation domain (Schön and Kutterer, 2005b).

4.2.4 Error bounding for combined uncertainties

An extended confidence region, denoted by $\mathcal{C}_E$, represents the combined measure of uncertainties due to both stochasticity and remaining systematics. In the context of set theory, this region is computed as the *Minkowski sum* of two convex sets (Schön and Kutterer, 2001a, 2006):

$$\begin{aligned}\mathcal{C}_E :=&\mathcal{E}_{1-\alpha}\oplus\mathcal{Z}_D\\ =&\Big\{\mathbf{x}\in\mathbb{R}^m\ \Big|\ (\mathbf{x}-\mathbf{z})^T\mathbf{Q}_{\hat{x}\hat{x}}^{-1}(\mathbf{x}-\mathbf{z})\le\chi^2_{1-\alpha,n},\\ &\text{with}\ \mathbf{z}=\hat{\mathbf{x}}+\sum_i^n w_i\cdot\mathrm{col}_i(\mathbf{K})\cdot\Delta_{s,i},\ -1\le w_i\le 1\Big\},\end{aligned} \tag{4.12}$$

where,

- $\mathcal{Z}_D$ represents the zonotope uncertainty due to remaining systematics;
- $\mathcal{E}_{1-\alpha}$ denotes the confidence ellipse with confidence level $1-\alpha$;
- $\mathcal{C}_E$ is the extended confidence region, which remains a symmetric convex set but no longer elliptical nor zonotopic.

The conceptual sketch in Fig. 4.1 provides a heuristic understanding of Eq. 4.12, illustrating how two classes of uncertainty interact. The point uncertainty due to random errors can be bounded by an elliptical range that is related to a confidence level. Each point within this range is also concerned with interval-bounded systematic errors. Therefore, the desired confidence region is defined by the envelope of zonotopes whose centers lie on the elliptical boundary. The resulting *extended confidence region* can be interpreted as follows:

> *Taking into account the uncertainty due to remaining systematics, the extended confidence region covers the true position with a probability of* $1-\alpha$ *for least-squares estimations* (Schön, 2003).

Let $\mathbf{x}^*$ denote the unknown true value of the state vector (e.g., receiver's position), and the estimation error bound is known by

$$\mathbf{x}^*\in\mathcal{C}_E \Leftrightarrow \|\varepsilon_{LS}\| := \|\mathbf{x}^*-\hat{\mathbf{x}}\|\le\mathrm{rad}(\mathcal{C}_E), \tag{4.13}$$

which is the radius of the extended confidence region, and the true position is assured to be covered with a lower bounded probability:

$$P\left(\mathbf{x}^*\in\mathcal{C}_E\right)=P\left(\|\varepsilon_{LS}\|\le\mathrm{rad}(\mathcal{C}_E)\right)\ge 1-\alpha \tag{4.14}$$

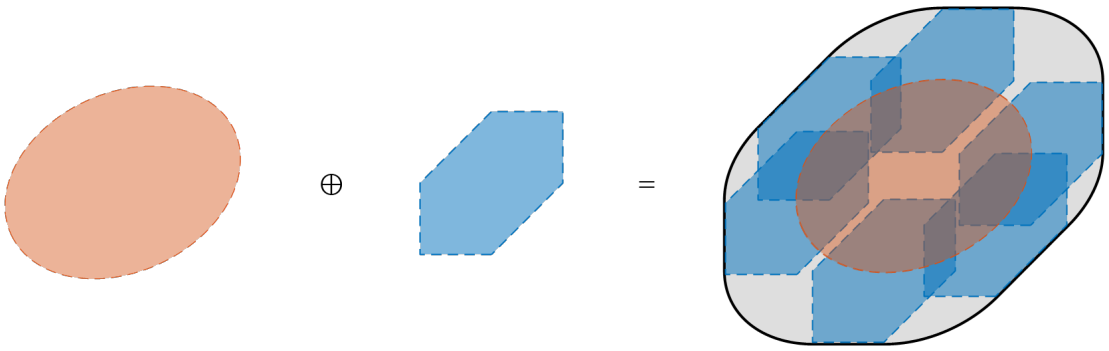


Figure 4.1: Minkowski sum of zonotope and confidence ellipse.

Eq. 4.12-4.13 are promising for GNSS integrity applications. They allow for addressing non-Gaussian error sources by distribution-free interval solutions instead of restricting them to Gaussian distributions when applying the least-squares estimator. This has the potential to reduce estimation error bounds (or the confidence region, equivalently) while ensuring the required confidence level.

In ARAIM, this approach can be directly utilized to calculate Protection Levels (PL). The PL is designed to accommodate not only stochastic nominal errors, which are overbounded by inflated Gaussian distributions, but also nominal biases, for which worst-case values are specified. In Sec. 6.1, the worst-case nominal bias is interpreted as interval bounds. This error bounding approach is applied to the ARAIM framework to reduce computed fault-free Horizontal Protection Level (HPL)s without increasing unacceptable complexities. The improvements with respect to the baseline ARAIM algorithm are then evaluated through a simulation analysis.

4.3 Set-based state estimation

The classical least-squares estimator makes use of all available measurements and provides an optimal solution in the least-squares sense, i.e., the ℓ_2 norm of residuals is minimum. However, it remains questionable whether the least-squares estimator is still representative or provides a "best" estimate for the unknown parameters in case of interval-described uncertainty of measurements. This section aims to fill the gap by (i) formulating the set estimator based on constraint satisfaction, (ii) proposing a central set-based estimator that is optimal in a deterministic sense, (iii) discussing the feasibility of the set-constrained least-squares estimator that takes advantage of the least-squares principle along with set constraint.

4.3.1 Set estimator based on constraint satisfaction

Estimation principle and concepts

Now consider again the measurement model in Eq. 4.2 but with different uncertainty handling. The VCM of the measurement error, i.e., Σ, is not accessible. Instead, it is known that the errors can be bounded by an interval vector $[\mathbf{e}] = [\underline{\mathbf{e}}, \overline{\mathbf{e}}]$, which is not necessarily symmetric. Subsequently, it is possible to characterize a *measurement uncertainty set (MUS)*:

> *The set including all possible measurement values whose errors will not exceed given interval bounds. The MUS contains all the information that could have generated the noisy measurements* (Casini, 2002).

Using $\mathbf{x}^*$ to denote the unknown true position of the GNSS receiver and neglecting the linearization error, the MUS reads:

$$[\mathbf{y}^*] \coloneqq \mathbf{A}\,\mathbf{x}^* + [\mathbf{e}]. \tag{4.15}$$

It is a subset of the $\mathbb{R}^n$ observation space in the form of an interval box (or interpreted as an n-dimensional hypercube).

To explore the state domain, the set membership can be formulated based on Eq. 4.2 as follows:

$$\mathbf{A}\,\mathbf{x} \in \mathbf{y} - [\mathbf{e}]. \tag{4.16}$$

Thanks to the definition of intervals, the following inequalities can be constructed:

$$\begin{cases} \mathbf{A}\,\mathbf{x} \leq \mathbf{y} - \underline{\mathbf{e}} \\ -\mathbf{A}\,\mathbf{x} \leq -\mathbf{y} + \overline{\mathbf{e}} \end{cases}. \tag{4.17}$$

Here, $\leq$ and $\geq$ denote component-wise inequalities between vectors (cf. the entry "$\leq$ and $\geq$" in the *Notations and Nomenclatures* section).

Applying $\mathbf{B} = \begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} \mathbf{y} - \underline{\mathbf{e}} \\ -\mathbf{y} + \overline{\mathbf{e}} \end{bmatrix}$ leads to a short version: $\mathbf{B}\,\mathbf{x} \leq \mathbf{b}$, which formulates an inequality-based constraint satisfaction problem. Solving this problem results in a solution set in the form of a H-polytope:

$$\mathcal{P}_s(\mathbf{B}, \mathbf{b}) = \left\{ \mathbf{x} \in \mathbb{R}^m \,|\, \mathbf{B}\,\mathbf{x} \leq \mathbf{b}, \mathbf{B} \in \mathbb{R}^{2n \times m}, \mathbf{b} \in \mathbb{R}^{2n} \right\}. \tag{4.18}$$

In an ideal situation, i.e., the expected measurements $\mathbf{y}^* \coloneqq \mathbf{A}\,\mathbf{x}^*$ do not present any random error nor systematic bias, the solution set is a H-zonotope centered at the true position ($\mathbf{x}^*$):

$$\mathcal{Z}_s(\mathbf{B}, \mathbf{b}^*) = \left\{ \mathbf{x} \in \mathbb{R}^m \;\middle|\; \mathbf{B}\,(\mathbf{x} - \mathbf{x}^*) \leq \mathbf{b}^*, \mathbf{B} \in \mathbb{R}^{2n \times m}, \mathbf{b}^* = \begin{bmatrix} -\underline{\mathbf{e}} \\ \overline{\mathbf{e}} \end{bmatrix} \in \mathbb{R}^{2n} \right\}. \tag{4.19}$$

$\mathcal{Z}_s$ is a set in the state domain that is directly mapped by the linearized model (cf. Eq. 4.2) from the MUS. Any candidate point of the zonotope can be traced back to an element in the MUS, i.e., a vector of measurement values $\mathbf{y} \in [\mathbf{y}^*]$. In this context, it can be interpreted as a *confidence region* for the desired set estimator.

If a set of measurement samples is given, the set solution becomes a reshaped and shifted polytope, deviating from the true position. The shape and position of the polytope $\mathcal{P}_s$ are jointly influenced by the measurement geometry, measurement samples, and interval bounds for the observation errors. Fortunately, it is still guaranteed that the true position is enclosed thanks to the linear propagation of interval-described uncertainty:

$$\mathbf{x}^* \in \mathcal{P}_s. \tag{4.20}$$

The elaboration is as follows. The interval bound $[e_i]$ represents the worst case that the error of i-th measurement can reach; hence, an admissible measurement value y_i would lie within the range $y_i^* + [e_i]$. Alternatively, given a measurement sample y_i, $y_i - [e_i]$ indicates the possible range of the expected measurement y_i^*. Subsequently, the interval box containing the expected $\mathbf{y}^*$, denoted by $[\mathbf{y}]$, is known immediately:

$$\begin{array}{c} \mathbf{y} \in [\mathbf{y}^*], \\ \text{with } [\mathbf{y}^*] \coloneqq \mathbf{y}^* + [\mathbf{e}]. \end{array} \quad \Rightarrow \quad \begin{array}{c} \mathbf{y}^* \in [\mathbf{y}], \\ \text{with } [\mathbf{y}] \coloneqq \mathbf{y} - [\mathbf{e}]. \end{array} \tag{4.21}$$

In this context, $\mathcal{P}_s$ can be understood as the result of mapping a set $[\mathbf{y}]$ from the $\mathbb{R}^n$ observation space to the $\mathbb{R}^m$ state space through the linearized model (Eq. 4.2). As such, Eq. 4.20 is proven.

Furthermore, it is reminded that the interval bounds are not necessarily tied to distributional assumptions, as discussed in Sec. 2.2. Therefore, no "exact" information is available regarding which point within $\mathcal{P}_s$ represents the truth, but all elements are treated equally. In this regard, $\mathcal{P}_s$ is referred to as a *feasible solution set (FSS)*:

> *The set consisting of all admissible solutions, compatible with the available information in the estimation problem* (Casini, 2002).

By definition, the worst-case estimation error of a set estimator equals the maximum possible Euclidean distance between any pair of points within the FSS:

$$\|\varepsilon_s\| \leq \max_{\mathbf{x}_1,\mathbf{x}_2 \in FSS} \|\mathbf{x}_1 - \mathbf{x}_2\|. \tag{4.22}$$

Geometrically, it is the *diameter* of the FSS, i.e.,

$$\|\varepsilon_s\| \leq \operatorname{diam}(FSS), \tag{4.23}$$

which represents a scalar measure of the estimation error bound for the set estimator.

It is important to note that some literature additionally defines the "feasible problem element set (FPS)". This term is not distinguished in this thesis since it aligns with the FSS concept in parameter estimation problems such as Eq. 4.2. Furthermore, the FSS can take on complex forms, e.g., highly irregular shapes. Hence, it may be beneficial to seek approximate representations for practical applications, such as using boxes or ellipsoids that either contain or are contained in the FSS (Milanese and Vicino, 1991; Schön, 2003).

In current GNSS applications, the standard practice is to provide a *point solution* along with an indication of estimation uncertainty rather than a *set solution*, although the latter would offer greater trustworthiness. Accordingly, the remainder of this chapter seeks to design alternative point estimators formulated as optimization problems based on the FSS for various purposes.

Extension to probabilistic framework

The set estimator can be extended to the probabilistic context. If probabilities are associated with interval bounds, the probability associated with the FSS can also be determined. For instance, P_Δ represents the probability that the i-th measurement error e_i does not exceed the interval range $[e_i] \coloneqq [\underline{e}_i, \overline{e}_i]$. Subsequently, the probability that Eq. 4.21 is valid for y_i is also P_Δ:

$$\begin{aligned} P\left(e_i \in [e_i]\right) =& P\left(y_i \in [y_i^* + \underline{e}_i, y_i^* + \overline{e}_i]\right) \\ =& P\left(y_i^* \in [y_i]\right) \\ =& P_\Delta, \end{aligned} \quad \text{with} \quad \begin{cases} y_i \coloneqq & y_i^* + e_i, \\ [y_i] \coloneqq & y_i - [e_i] \\ = & [y_i - \overline{e}_i, y_i - \underline{e}_i]. \end{cases} \tag{4.24}$$

In the remainder of the thesis, this probability is termed *bounding probability*. According to the binominal law, the probability that Eq. 4.21 is valid for all *independent* measurements simultaneously is (Drevelle and Bonnifait, 2011)

$$\begin{aligned} P\left(\mathbf{y}^* \in [\mathbf{y}]\right) \coloneqq & P\left(y_1^* \in [y_1] \cap \ldots \cap y_n^* \in [y_n]\right) \\ =& P\left(y_1^* \in [y_1]\right) \cdot \ldots \cdot P\left(y_n^* \in [y_n]\right) \\ =& \left(P_\Delta\right)^n, \end{aligned} \tag{4.25}$$

which equals $P\left(\mathbf{x}^* \in \mathcal{P}_s\right)$ due to linear mapping, i.e., the probability that the set solution described in Eq. 4.18 encloses the truth.

Taking advantage of this feature, as an example, we can apply the set estimator under the Gaussian-related assumptions, which is widely used in GNSS applications. Precisely, the observation errors are assumed to be independent Gaussian distributed or can be overbounded by a Gaussian distribution $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \Sigma)$, with Σ containing the variance of each measurement error σ_i^2. In this context, seeking a worst-case error bound becomes meaningless due to the infinite tails of the distributions. Nevertheless, the bounding probability can be calculated based on the required confidence level $1 - \alpha$ for estimation.

Let $[e_i]$ be an interval that is symmetric around zero for the i-th measurement error; its radius Δ_i is calculated as a product of a coefficient k_α and the standard deviation of the measurement error σ_i:

$$[e_i] = [-\Delta_i, \Delta_i], \ \Delta_i = k_\alpha \cdot \sigma_i. \tag{4.26}$$

The coefficient k_α can be determined by

$$k_\alpha = Q^{-1}\left(\frac{P_\Delta}{2}\right), \tag{4.27}$$

$$\text{with } (P_\Delta)^n \geq 1 - \alpha. \tag{4.28}$$

In Eq. 4.28, the product $(P_\Delta)^n$ is lower bounded by the confidence requirement $1 - \alpha$, which can be used to determine P_Δ. Then, k_α is computed using the assumed Gaussian distribution in Eq. 4.27, where $Q^{-1}(P)$ is the $(1 - P)$ quantile of the standard normal distribution.

The aforementioned strategy opens the possibility of handling stochastic error components by constructing probability-associated interval bounds and applying them to the set estimator. This aspect is further explored in the comparison analyses and application examples in the remainder of the thesis.

Demonstrative example

A two-dimensional positioning example is designed to demonstrate the performance of the set estimator, utilizing the following parameters:

- Measurement vector: $\mathbf{y} = [0.5083, 2.7495, -2.5043, -0.2377]^T$, i.e., four observations $n = 4$,
- Geometry matrix: $\mathbf{A} = \begin{bmatrix} 0.2588 & 0.9659 \\ 0.5000 & 0.8660 \\ 0.7071 & -0.7071 \\ 0.9659 & -0.2588 \end{bmatrix}$,
- True states (the coordinate increments): $\mathbf{x}^* = [0, 0]^T$, i.e., the initial guess coincides with the true position, so that the expected observations $\mathbf{y}^* = \mathbf{A}\,\mathbf{x}^* = [0, 0, 0, 0]^T$.

At first, three example interval bounds are specified for measurement errors ($[-\Delta_i, \Delta_i]$ with the radius Δ_i). These bounds are determined from some uncertainty model, which remains unspecified at this stage. Applying these intervals to the set estimator leads to different polytope solutions. Comparing these intervals with the simulated measurement errors culminates in different scenarios concerning measurement error bounding:

1. **Adequate bounding.** The first interval radius $\Delta_i = 4.708$ adequately bounds all simulated measurement errors. The resulting polytope set solution is illustrated as the green contour in Fig. 4.2, enclosing the truth marked by a red star.
2. **Tightened bounding.** Reducing the interval radius to 3.662 demonstrates the situation with tightened measurement error bounds (e.g. when more information on measure-

ment uncertainty is available). This adjustment results in a smaller, yet still sufficient, polytope solution depicted as the yellow contour in Fig. 4.2.

3. **Inadequate bounding.** When applying the smallest radius ($\Delta_i = 2.226$), the third measurement error exceeds the interval bound. The polytope solution under this setting is shown in blue in Fig. 4.2, failing to encompass the truth.

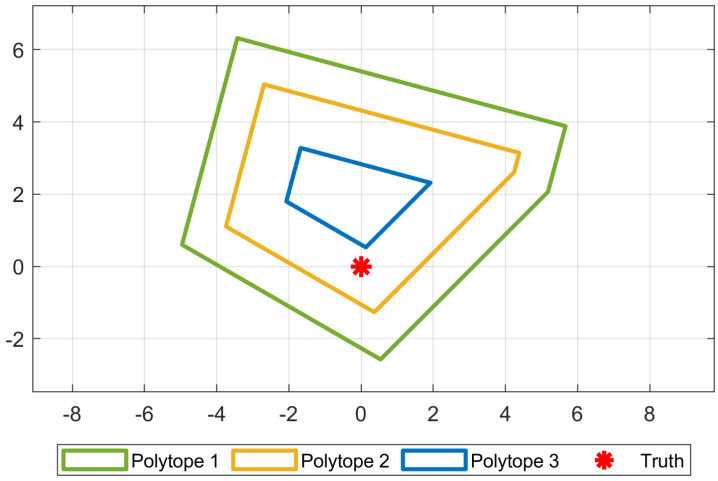


Figure 4.2: A numerical demonstrative example for the set estimator, with solutions obtained using different interval bounds.

Alternatively, this example can be interpreted from a stochastic perspective. If the prior distribution of the measurement error is available, i.e., $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \Sigma = \mathbf{I})$, interval bounds should be probabilistically defined based on the desired confidence level α, cf. Eq. 4.26 to Eq. 4.28. Three example confidence levels are shown in Tab. 4.2, resulting in various interval radii. Correspondingly, Fig. 4.2 implies how the set solution varies with the confidence level, i.e., lower bounding probabilities increase the chance of the set solution failing to cover the truth.

Table 4.2: Confidence levels, bounding probabilities, and interval radii in the demonstrative example.

Parameter	**Set solution**		
	Polytope 1	Polytope 2	Polytope 3
Confidence level $1-\alpha$	$1-10^{-5}$	$1-10^{-3}$	$1-10^{-1}$
Bounding probability P_Δ	$1-2.5\times10^{-6}$	$1-2.5\times10^{-4}$	$1-2.6\times10^{-2}$
Interval radius $\Delta_i = k_\alpha\,\sigma_i$	4.708	3.662	2.226

4.3.2 Optimal set-based central estimator

Deterministic optimality concept

Set-membership estimation theory incorporates concepts of optimality borrowed from information-based complexity theory (Traub et al., 1988). Referring to the sets defined in Sec. 4.3.1, three types of optimality for an estimator can be identified. Let $\hat{\mathbf{x}} = g(\mathbf{y})$ denote the desired estimator for the problem described in Eq. 4.2, which is a function of the measurement vector, and ε denote the estimation error. The three types of optimality can then be formulated as follows:

- **X-locally optimal.** Let $\mathbf{x} \in \mathbb{R}^m$, the *X-local error*

$$\varepsilon_X \coloneqq \max_{\mathbf{y} \in MUS} \|\mathbf{x} - \hat{\mathbf{x}}\|, \quad \text{with } \hat{\mathbf{x}} = g(\mathbf{y}) \tag{4.29}$$

  is minimized when the measurements are not known, and $\hat{\mathbf{x}}$ offers the best estimate related to the worst value of $\mathbf{y}$ in MUS. The X-local error can be considered as an *a priori* measure since it is not based on the observations.

- **Y-locally optimal.** Let $\mathbf{y} \in \mathbb{R}^n$, the *Y-local error*

$$\varepsilon_Y \coloneqq \max_{\mathbf{x} \in FSS} \|\mathbf{x} - \hat{\mathbf{x}}\|, \quad \text{with } \hat{\mathbf{x}} = g(\mathbf{y}) \tag{4.30}$$

  is minimized when a set of sample values for the measurement $\mathbf{y}$ is available and, hence, $\hat{\mathbf{x}}$ is the best achievable estimate for $\mathbf{y}$. Different from the X-local error, the Y-local error reflects the *posteriori* goodness of the estimation. Notably, a necessity for the Y-local optimality is that FSS is non-empty; otherwise, the above expression becomes meaningless. For example, the FSS derived in Sec. 4.3.1 is guaranteed to be non-empty, providing *worst-case* measurement intervals (at least $\mathbf{x}^* \in \mathcal{P}_s$).

- **Globally optimal.** The *global error*

$$\varepsilon_G \coloneqq \max_{\mathbf{x} \in \mathbb{R}^m} \varepsilon_X = \max_{\mathbf{y} \in \mathbb{R}^n} \varepsilon_Y \tag{4.31}$$

  is minimized. In this regard, two possibilities exist for achieving global optimality because of the dual definitions of global error.

It follows that the choice of a favorite "optimal" estimator depends on the condition of the investigated estimation problem. Given measurement samples $\mathbf{y}$, the Y-local optimality is desired where the worst-case error is minimized for all observed values. On the contrary, the X-locally optimality gives the best estimation related to the worst possible value of measurements despite the absence of sample values. Furthermore, it is noticed that the local optimality is stronger than global optimality (Milanese and Vicino, 1991). For instance, an estimator is not only Y-locally optimal but also provides the minimal "worst-case" Y-local error for all possible measurement samples $\mathbf{y} \in \mathbb{R}^n$, then, it is also globally optimal. However, the converse is not necessarily true.

Considering the introduced optimality conditions, an optimal estimator for the model in Eq. 4.2 is expected to account for the following scenario:

> *Y-locally optimal, therefore, also globally optimal, provided interval-described uncertainty for the measurements.*

The measurement errors are known to be bounded by interval values, regardless of distributions within the range. Hence, they are interpreted as worst-case variations of the measurement errors. In this context, the resulting point solution is not linked with any information about the error distribution and is concerned with the worst-case situation.

Seeking the Y-local optimality is straightforward and directly driven by the problem context. For instance, observation error sources are primarily evaluated through interval-described uncertainty models. This evaluation allows for the use of the set estimator based on constraint satisfaction. Given a set of measurement samples, the outcome of the set estimator (regarded as a FSS in Eq. 4.18) then serves as *a priori* information to derive a "best" point solution. According to the deterministic optimality concepts, the desired estimator is naturally Y-locally optimal and, as a consequence, globally optimal.

Optimal central estimator

Designing the optimal point estimator starts from the following assumption:

$$\mathbf{x}^* \in FSS, \tag{4.32}$$

which is fulfilled conditioning on adequate interval bounding for measurement errors. For example, Eq. 4.20 shows the case where the outcome of the set estimator shall be used.

Denoting the desired estimator as $\hat{\mathbf{x}}_c$, the worst-case error in the state domain can be expressed as

$$\|\varepsilon_c\| \coloneqq \|\mathbf{x}^* - \hat{\mathbf{x}}_c\| \leq \max_{\mathbf{x}^* \in FSS} \|\mathbf{x}^* - \hat{\mathbf{x}}_c\|. \tag{4.33}$$

It follows that $\hat{\mathbf{x}}_c$ yields Y-locally optimality and, naturally, global optimality:

$$\hat{\mathbf{x}}_c = \arg \min_{\mathbf{x} \in \mathbb{R}^m} \max_{\mathbf{x}^* \in FSS} \|\mathbf{x}^* - \mathbf{x}\|. \tag{4.34}$$

The interpretation is as follows. Considering all candidates of truth being within FSS, $\hat{\mathbf{x}}_c$ leads to the minimum-achievable worst-case estimation error. It is the *Chebyshev center* of FSS by definition (cf. Eq. 2.8):

$$\hat{\mathbf{x}}_c = \text{cen}(FSS), \tag{4.35}$$

and the corresponding worst-case error can be interpreted as the *Chebyshev radius* of FSS (cf. Eq. 2.9), i.e., $\max_{\mathbf{x}^* \in FSS} \|\mathbf{x}^* - \hat{\mathbf{x}}_c\| = \text{rad}(FSS)$. Thereby, the estimation error of the desired central estimator can be upper bounded by

$$\|\varepsilon_c\| \leq \text{rad}(FSS). \tag{4.36}$$

It is noteworthy that the proposed estimator does not rely on any assumptions regarding the distributions of measurement errors. Instead, it depends entirely on the *a priori* FSS, which captures the uncertainty in the state domain propagated from interval-described uncertainties at the measurement level. This approach ensures optimality in minimizing the worst-case estimation error. The shape of the FSS defines the estimation error bounds, while the Chebyshev radius of FSS provides a scalar measure of these bounds.

The following properties of the central estimator regarding the convexity of FSS are observed:

- If FSS forms a convex set, Eq. 4.34 constitutes a convex optimization problem where an optimal and unique solution is achievable.
- If FSS is non-convex, the point solution might not reside within the FSS.

Thus, the set-based central estimator is particularly practical in the case of convex FSS, for instance, when using the outcome of the set estimator based on constraint satisfaction. If the scenario is different, the convexity of the FSS should be carefully examined.

Suboptimal central estimator

Finding the Chebyshev center of a set may be a difficult and typically intractable problem, with two exceptional cases (Eldar et al., 2008):

- When the set is polyhedral;
- When the set is finite.

Fortunately, the set solution $\mathcal{P}_s$ from the set estimator based on constraint satisfaction, cf. Eq. 4.18, satisfies these requirements and, hence, can provide an optimal point solution.

In certain scenarios, a suboptimal estimator may be of particular interest for practical reasons. For example, real-time processing often requires high computational efficiency, which can be limited by the capabilities of onboard hardware. Therefore, instead of computing the Chebyshev center of the exact FSS, it is preferable to seek the center of the outer approximation of FSS, denoted by $\overline{\mathcal{S}}$, such that $FSS \subseteq \overline{\mathcal{S}}$. The desired outer approximation remains convex and, hence, is applicable to Eq. 4.34:

$$\hat{\mathbf{x}}'_c \coloneqq \arg \min_{\mathbf{x}\in\mathbb{R}^m} \max_{\mathbf{x}^*\in\overline{\mathcal{S}}} \|\mathbf{x}^* - \mathbf{x}\| = \text{cen}(\overline{\mathcal{S}}). \tag{4.37}$$

Some properties of $\hat{\mathbf{x}}'_c$ may be noticed:

- It is contained in $\overline{\mathcal{S}}$;
- It does not necessarily coincide with the optimal central estimator $\hat{\mathbf{x}}_\mathbf{c}$;

Furthermore, if FSS is associated with a probability $P(\mathbf{x}^* \in FSS)$, i.e., the probability that the truth is covered by FSS, it is easy to show that the probability associated with the outer approximated set is no smaller than that associated with FSS:

$$FSS \subseteq \overline{\mathcal{S}} \Rightarrow P(\mathbf{x}^* \in FSS) \leq P(\mathbf{x}^* \in \overline{\mathcal{S}}). \tag{4.38}$$

In complex systems, the set estimator based on constraint satisfaction may not be directly usable, or multiple set solutions must be estimated, contributing to the final FSS. Subsequently, the obtained FSS may be non-convex. To address this, the exact non-convex FSS in Eq. 4.34 can be replaced with its outer convex approximation to make the problem solvable.

The outer approximation can take various forms, such as:

- The convex hull of FSS for non-convex FSS: $\text{conv}(FSS)$.
- The axis-aligned interval bounding box containing FSS: $[\mathbf{x}_s]$, referred to as *minimum volume outer box* in literature such as Milanese and Vicino (1991).
- Other regular-shaped sets that circumscribe FSS. For instance, researchers have investigated ellipsoids (Schweppe, 1968; Kurzhanskiy and Varaiya, 2007), zonotopes (Le et al., 2013; Scott et al., 2016), parallelotopes (Chisci et al., 1996), etc., in various problems.

Specifically, in the case of axis-aligned interval bounding box $\overline{\mathcal{S}} = [\mathbf{x}_s]$, the point solution can be computed state-wise as the center of each interval component of $[\mathbf{x}_s]$:

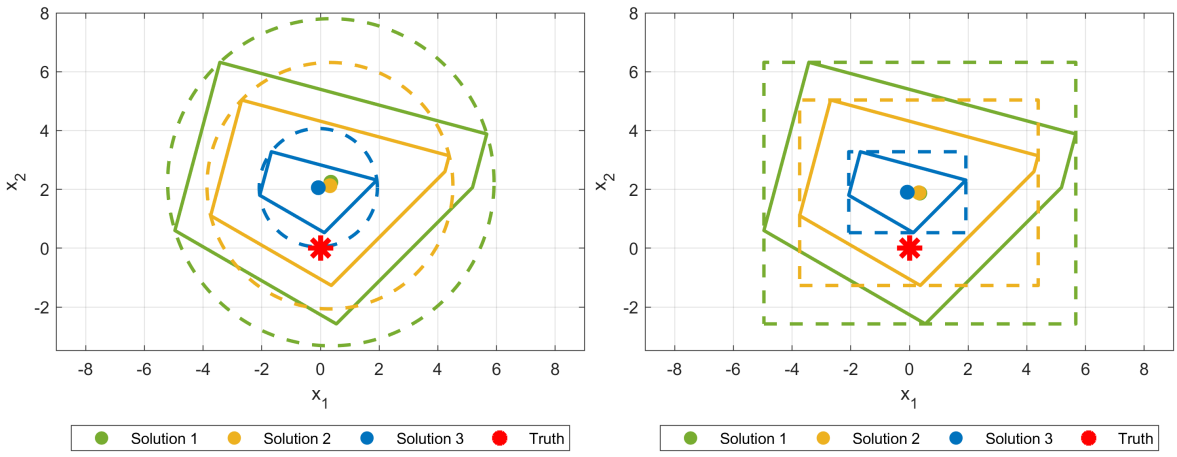
$$\begin{aligned} \hat{x}'_{c,i} &\coloneqq \arg \min_{\mathbf{x}\in\mathbb{R}^m} \max_{\mathbf{x}^*\in\overline{\mathcal{S}}} \|x_i^* - x_i\| \\ &= \text{cen}([x_{s,i}]) \\ &= \text{cen}\left(\left[\min_{\mathbf{x}\in FSS} x_i, \max_{\mathbf{x}\in FSS} x_i\right]\right) \text{ for } i = 1, ..., m, \end{aligned} \tag{4.39}$$

where the subscript i represents the i-th component of the variable.

Demonstrative example

The two-dimensional positioning example from Sec. 4.3.1 is revisited to elucidate the features of the set-based central estimator.

The FSS is determined using the set estimator based on constraint satisfaction, and the optimal central estimator computes for a point solution as the Chebyshev center of FSS. In Fig. 4.3(a), various point solutions (shown as colored dots) are obtained for varying interval radii, and the corresponding metrics are given in Tab. 4.3 (i.e., estimation errors and error bounds). Notably, when the smallest polytope fails to encompass the truth, the worst-case error bound does not upper bound the position error, emphasizing the importance of adequate interval bounding for measurement errors.



(a) Result of optimal set-based central estimator (b) Result of suboptimal set-based central estimator

Figure 4.3: A demonstrative example for the set-based central estimator. The solutions are obtained using different observation interval bounds.

Table 4.3: Numerical results of the set-based central estimators in the demonstrative example. Three feasible solution sets, determined using different interval bounds, are fed into the optimal and suboptimal set-based central estimators, corresponding to various confidence levels under the Gaussian assumption.

Parameter	**Feasible solution set**		
	Polytope 1	Polytope 2	Polytope 3
Interval radius Δ	4.708	3.662	2.226
Optimal central estimator			
Estimation error $\|\varepsilon\|$	2.270	2.149	2.059
Worst-case error bound	5.561	4.190	2.013
Suboptimal central estimator (state-wise)			
Estimation error $(\|\varepsilon_1\|, \|\varepsilon_2\|)$	$0.355, 1.872$	$0.323, 1.885$	$0.073, 1.903$
Worst-case error bound (x_1, x_2)	$5.314, 4.447$	$4.064, 3.153$	$1.996, 1.376$
Confidence level $1-\alpha$	$1-10^{-5}$	$1-10^{-3}$	$1-10^{-1}$

In scenarios where the interval bounds are set unrealistically narrow, the FSS might become empty, preventing the derivation of a point solution from the central estimator. This is a different behavior from the classical least-squares estimator that will, in all cases, lead to a point solution. Unlike the classical least-squares estimator, which can, in all cases, provide a point solution, the set-based central estimator exhibits robustness against outliers since the estimation will not be biased severely.

Re-consider this example within the probabilistic framework (as in Tab. 4.2). It is observed that the smallest polytope reflects the lowest confidence level $1-\alpha$ with the worst bounding probabilities for interval bounds and vice versa. This observation can be interpreted as follows:

The higher the bounding probabilities of the interval bounds are, the more confidence we have when taking the Chebyshev radius of FSS as the estimation error bound for the set-based central estimator.

Suboptimal point solutions can be computed from the axis-aligned bounding box containing FSS, cf. Eq. 4.39, shown in Fig. 4.3(b). The estimation errors and error bounds are summarized in Tab. 4.3, which are slightly different from optimal solutions. Although it appears, in this case, that the suboptimal solutions are more accurate, the tendency is not conclusive and can vary from time to time. Nevertheless, the suboptimal central estimator is advantageous in computational efficiency as it requires only a sorting function rather than solving a two-dimensional Chebyshev center problem.

4.3.3 Set-constrained least-squares estimator

Introduction

Consider a scenario where measurement errors are known to be bounded by interval values, which may or may not be associated with probabilities. Additionally, an "overbounding" variance is provided. A representative example is the residual tropospheric error. Traditional integrity methods treat such errors stochastically, assuming that they are overbounded by Gaussian distributions with inflated variances that are computed from extensive historical datasets. In contrast, Sec. 3.2 presents an adequate interval bounding approach, supported by experimental analysis.

Within this context, the objective is to develop an estimator that yields the "best" point solution by incorporating the following aspects:

- Assurance of componentwise interval bounds for measurement errors.
- Retention of the least-squares principle.
- Availability of a sufficiently accurate initial position estimate to ensure valid linearization.

These factors motivate the formulation of the *set-constrained least-squares (SCLS)* estimator.

Considering the model in Eq. 4.2 and Eq. 4.17, the desired estimator can be formulated as:

$$\begin{aligned} &\text{minimize } (\mathbf{y} - \mathbf{A}\,\mathbf{x})^T \Sigma^{-1} (\mathbf{y} - \mathbf{A}\,\mathbf{x}), \\ &\text{subject to } \mathbf{x} \in FSS, \end{aligned} \tag{4.40}$$

where the feasible solution set contains all admissible solutions compatible with interval error bounding for observations, hence forming a closed constrained region (i.e., a set constraint).

When determining the set constraint imposed by the set estimator based on constraint satisfaction, the SCLS estimator is a special class of the *inequality-constrained least-squares (ICLS)* estimator. The problem can be formulated as follows:

$$\begin{aligned} &\text{minimize } (\mathbf{y} - \mathbf{A}\,\mathbf{x})^T \Sigma^{-1} (\mathbf{y} - \mathbf{A}\,\mathbf{x}), \\ &\text{subject to } \mathbf{B}^T\mathbf{x} \leq \mathbf{b}, \quad \text{with } \mathbf{B} = \begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} \mathbf{y} - \underline{\mathbf{e}} \\ -\mathbf{y} + \overline{\mathbf{e}} \end{bmatrix}. \end{aligned} \tag{4.41}$$

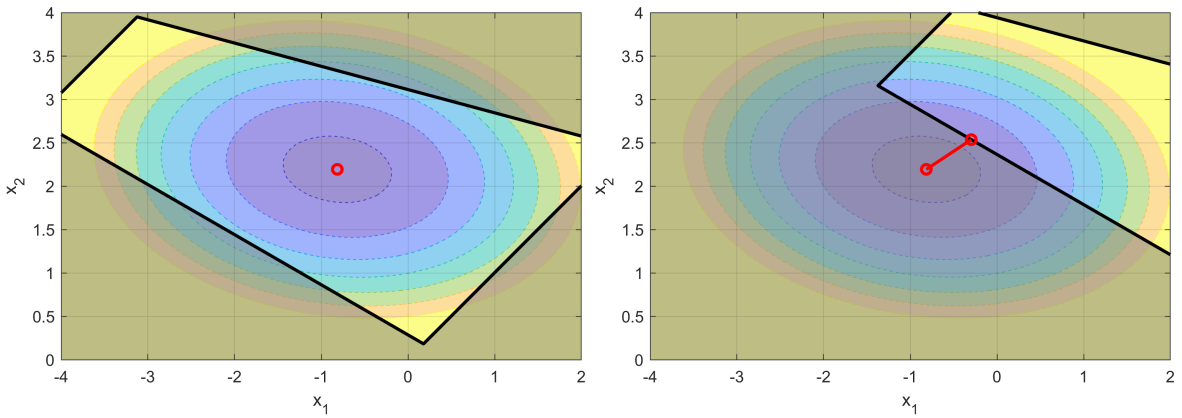
The polytope set constraint ($\mathcal{P}_s$) is mathematically equivalent to these inequality constraints. Therefore, any method used to solve a general ICLS problem is applicable to Eq. 4.41. The remainder of this section focuses on Eq. 4.41 to maintain consistency and avoid introducing unnecessary complexity in the context.

Solving the SCLS problem

By definition, the SCLS problem is solved in two steps:

1. Defining the objective function (the weighted sum of squared residuals). In Fig. 4.4, this objective function is represented by colored elliptical isolines centered around the classical least-squares solution.

2. Transferring measurement-level intervals into the position domain linearly as a "hard" set constraint. The truth is presumed to reside in the set constraint (FSS) with a certain confidence, which is depicted by the black contour in Fig. 4.4.



(a) Example of inactive set constraint for SCLS

(b) Example of active set constraint for SCLS

Figure 4.4: Conceptual sketch for an SCLS problem. Isolines of the objective function are plotted as colored elliptical contours and set constraints as black polygons. The classical least-squares solution is enclosed by the set constraint in Fig. 4.4(a), i.e., none of the inequality constraints is active. Hence, the resulting SCLS solution is invariant from the least-squares solution. One constraint appears active in Fig. 4.4(b) as the least-squares solution is located in the infeasible region. The SCLS solution is the projection of the unconstrained solution onto the boundary of the set constraint (two connected red dots).

The optimal solution for the given SCLS problem is then the point within the set constraint that minimizes the objective function. If the classical least-squares solution is contained within the set constraint, it is accepted as the SCLS solution (Fig. 4.4(a)). If any inequality constraints are active, the classical least-squares solution lies outside the set constraint, necessitating the search for an alternative point that is typically located at the boundary of the set constraint (Fig. 4.4(b)).

The general ICLS estimator and consequently the desired SCLS can be regarded as convex optimization problems where a quadratic function (the weighted sum of squared residuals) is to be minimized subject to linear inequality constraints (Eq. 4.40) (Boyd and Vandenberghe, 2004). Techniques suitable for solving such problems include the simplex methods (e.g., active set method) and interior point methods.

Several methods have been investigated to address specific questions that can be described as ICLS problems. For example, Liew (1976) focused on identifying active constraints to convert the ICLS estimator to an equality-constrained one; Geweke (1986) advocated a Bayesian approach, which was later on extended and applied to GPS positioning with height constraints by Zhu et al. (2005); Werner (1990) utilized the theory of generalized inversion and developed a method yielding closed-form solutions; Peng et al. (2006) suggested merging multiple simple inequality constraints into a single complex equality constraint in a least-squares problem; Mead and Renaut (2010) formulated the box constraints as quadratic constraints, establishing an unconstrained regularized least-squares problem applicable to hydrological problems; Yuan

et al. (2020) adapted the ICLS problem into a corresponding linear complementarity problem for estimating LEO-GPS receiver differential code bias (DCB).

However, challenges persist in uncertainty qualification due to the statistical limitations of the estimator imposed by inequalities (Werner, 1990), which is crucial for their usage in engineering applications. For example, the VCM of the SCLS estimation is assessed using the converted equality-constrained least-squares estimator in Liew (1976); O'Leary and Rust (1986) examined the confidence interval estimation for a least-squares estimator under a non-negative constraint; Zhu et al. (2005) explored the posterior distribution and the mean square error (MSE) matrix of the parameters; Roese-Koerner et al. (2012) proposed a framework for the quality description of SCLS estimates via the Monte Carlo method, emphasizing that accumulation of densities at the constrained boundaries is more realistic to account for the impact of constraints. However, there are two main concerns with respect to existing approaches:

- Potentially unrealistic uncertainty modeling. For example, the treatment of the truncated probability mass may lead to over-optimistic results;
- Potential unsuitability to the set-constrained estimation problem. For instance, constraints are oftentimes fixed regardless of observation geometry and uncertainty, which is not in line with the current problem of interest.

To overcome these limitations, the confidence region is calculated using set theory, rather than focusing on the exact PDF. This is further elaborated in the following subsection.

Constrained confidence region

For a standard least-squares estimator, the uncertainty due to stochastic errors can be modeled as a confidence ellipse $\mathcal{E}_{1-\alpha}$ associated with a confidence level $1-\alpha$. Taking additionally the set constraint ($\mathcal{P}_s$) into account, we get a constrained confidence region for the SCLS estimator as the intersection of the two sets:

$$\mathcal{C}_c = \mathcal{E}_{1-\alpha} \cap \mathcal{P}_s, \tag{4.42}$$

which can be interpreted as the possible range of the unknown truth at the required probability. The following arguments support its validity:

- The definition of SCLS is free of any distributional assumptions about the actual errors.
- Overbounding of the actual errors is deemed achievable using both approaches, i.e., stochastic and interval-based error bounding.

While implementing the second argument may rely on assumptions related to error distributions (e.g., the overbounding variance, cf. Sec. 2.4.2), the obtained constrained confidence region can be simply seen as the interaction of two estimators, namely, the least-squares estimator and set estimator. Their estimation error bounds compensate for each other, leading to a contracted confidence region that yields the required probability of safety:

$$\begin{aligned} P(\mathbf{x}^* \notin \mathcal{C}_C) \leq & P(\mathbf{x}^* \notin \mathcal{P}_s \vee \mathbf{x}^* \notin \mathcal{E}_{1-\alpha}) \\ \leq & P(\mathbf{x}^* \notin \mathcal{P}_s) + P(\mathbf{x}^* \notin \mathcal{E}_{1-\alpha}) - \\ & P(\mathbf{x}^* \notin \mathcal{P}_s) \cdot P(\mathbf{x}^* \notin \mathcal{E}_{1-\alpha}) \\ \leq & \alpha_p + \alpha_e, \end{aligned} \qquad \text{with} \begin{cases} P(\mathbf{x}^* \notin \mathcal{P}_s) \leq \alpha_p, \\ P(\mathbf{x}^* \notin \mathcal{E}_{1-\alpha}) \leq \alpha_e, \end{cases} \tag{4.43}$$

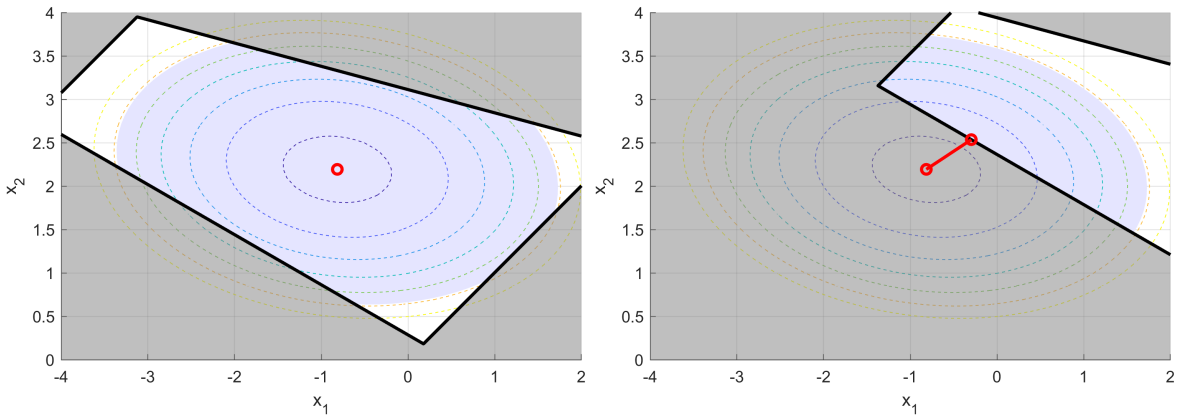
which implies that the probability of $\mathcal{C}_C$ not covering the truth is no greater than $\alpha = \alpha_p + \alpha_e$.

This inequality of probabilities eliminates the need to examine the independence of the two events ($\mathbf{x}^* \notin \mathcal{P}_s$ and $\mathbf{x}^* \notin \mathcal{E}_{1-\alpha}$), despite the potential conservatism ($P(\mathbf{x}^* \notin \mathcal{P}_s) \cdot P(\mathbf{x}^* \notin \mathcal{E}_{1-\alpha})$

is neglected). Subsequently, a tradeoff should be considered concerning the allocation of confidence requirements, i.e., α_p and α_e, which would highly affect the shape of the resulting constrained confidence region. This feature is further explored in the comparative analysis in Sec. 4.4.3.

A highly suitable scenario for applying the SCLS estimator is when deterministic interval error bounds are available. In this case, the intervals do not necessarily have any probabilistic interpretations. Consequently, Eq. 4.43 can be simplified to the following form:

$$P(\mathbf{x}^* \notin \mathcal{E}_{1-\alpha}) \leq \alpha_e \Rightarrow P(\mathbf{x}^* \notin \mathcal{C}_C) \leq \alpha_e. \tag{4.44}$$



(a) Constrained confidence region for Fig. 4.4(a) (b) Constrained confidence region for Fig. 4.4(b)

Figure 4.5: Conceptual sketch of the SCLS constrained confidence region for the examples in Fig. 4.5: The confidence ellipses are shown as colored dashed contours corresponding to various confidence levels, centered at the least-squares solution, and set constraints as the black polygons. The constrained confidence regions are represented by the intersection of a dedicated ellipse and the polytope, depicted as light blue areas. In both cases, the set constraint is active when determining the constrained confidence region.

Fig.4.5 are conceptual sketches of the constrained confidence region for Fig. 4.4. The polytope set constraint is obtained by the set estimator in black contours. In Fig. 4.5(a), the standard least-squares solution is depicted in red, which overlaps with the SCLS solution because the set constraint appears inactive when determining the point solution. Meanwhile, the SCLS solution differs from the least-squares solution in Fig. 4.5(b) due to the active set constraint. In both cases, the set constraint is active in determining the constrained confidence region, which remains a convex set and is no greater than either the polytope or the ellipse.

With the constrained confidence region $\mathcal{C}_C$, a scalar measure of the estimation error bound can be defined as the maximum possible Euclidean distance between the estimated point ($\hat{\mathbf{x}}$) and any admissible point of the truth ($\mathbf{x}$). This can be expressed mathematically as:

$$\|\varepsilon_{SC}\| \leq \max_{\mathbf{x} \in \mathcal{C}_\mathbf{C}} \|\hat{\mathbf{x}} - \mathbf{x}\|, \tag{4.45}$$

indicating the worst-case estimation error at a specified confidence level. It is noticed that

$$\mathrm{rad}(\mathcal{C}_C) \leq \|\varepsilon_{SC}\| \leq \mathrm{diam}(\mathcal{C}_C). \tag{4.46}$$

This means that the worst-case estimation error of the SCLS estimator is lower bounded by the radius of the constrained confidence region while upper bounded by its diameter.

4.4 Comparitive analysis of different estimators

This section aims to compare the investigated estimators from different aspects. The analysis begins with a qualitative comparison of their characteristics to provide a comprehensive insight into these estimators, followed by a theoretical evaluation for a simplified one-dimensional problem. Finally, a two-dimensional positioning problem is designed to evaluate their performance through statistical analysis, focusing on estimation accuracy and error bounding.

4.4.1 Qualitative comparison

Tab. 4.4 provides a detailed overview of the characteristics of the classical least-squares estimator, set estimator, set-based central estimator, and set-constrained least-squares estimator:

Confidence region. All estimators are capable of addressing both stochastic and systematic uncertainties, thus defining corresponding confidence regions. The least-squares estimator accounts for stochasticity with a confidence ellipse, and the remaining systematics are addressed by the interval extension of the least-squares estimator; thus an extended confidence region is deemed adequate. The set, central, and SCLS estimators define their confidence regions via interval mathematics and set theory, with measurement-level interval bounds fundamental to their operation.

Optimality. The least-squares estimator is optimal under the best linear unbiased estimation (BLUE) principle, offering the least variance of the estimate. If the variance of observations is only known with an upper-bounded value and interval bounds are available, the SCLS estimator can provide a modified least-squares solution that guarantees a minimally acceptable performance. This modification enhances the robustness of the least-squares estimator. In contrast, the set-based central estimator is optimal in terms of estimation error bounding by leveraging the set estimator based on constraint satisfaction.

Estimation error bounding. For both the least-squares estimator and set-based central estimator, the estimation error bound can be simply determined by the Chebyshev radius of their confidence regions. For the set estimator, this is the diameter of the polytope set solution, and for the SCLS estimator, it is the maximum Euclidean distance between the estimated point and the boundary of the confidence region.

4.4.2 Theoretical evaluation for a one-dimensional problem

To better understand the estimators' behavior, a simplified one-dimensional scenario based on the linearized model in Eq. 4.2 is considered. In this case:

- The geometry matrix (vector of $1 \times n$) $\mathbf{A} = [1 \ ... \ 1]$;
- $\mathbf{y}^* = \mathbf{A}\,\mathbf{x}^* = \mathbf{0}$, i.e., the unknown truth is assumed to be at the origin, and the initial guess of the state is overlapped with the truth;
- Observation errors $\mathbf{e}$ follow independent and identical continuous distributions with variances $\sigma_i^2 = 1$, i.e., $\Sigma = \mathbf{P}^{-1} = \mathbf{I}^n$. Accordingly, identical interval bounds $[\Delta_s, \Delta_s]$ are defined and assigned to observations.

In this case, each estimator behaves as follows:

Table 4.4: Comparison of the properties of different estimators and associated uncertainty modeling.

Property	**Estimator**			
	Least-squares estimator	Set estimator based on constraint satisfaction	Set-based central estimator[3]	Set-constrained least-squares estimator (SCLS)[3]
Estimation principle	Minimize $(\mathbf{y}-\mathbf{A}\,\mathbf{x})^T\Sigma^{-1}(\mathbf{y}-\mathbf{A}\,\mathbf{x})$	$\mathbf{B}\,\mathbf{x}\leq\mathbf{b}$, with $\mathbf{B}=\begin{bmatrix}\mathbf{A}\\-\mathbf{A}\end{bmatrix}, \mathbf{b}=\begin{bmatrix}\mathbf{y}-\underline{\mathbf{e}}\\-\mathbf{y}+\overline{\mathbf{e}}\end{bmatrix}$	Minimize $\max_{\mathbf{x}^*\in FSS}\|\mathbf{x}^*-\mathbf{x}\|$	Minimize $(\mathbf{y}-\mathbf{A}\,\mathbf{x})^T\Sigma^{-1}(\mathbf{y}-\mathbf{A}\,\mathbf{x})$, subject to $\mathbf{x}\in FSS$
Solution type	Point	Set	Point	Point
Estimation solution	$\hat{\mathbf{x}}=(\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{P}\mathbf{y}$	$\mathcal{P}_s=\{\,\mathbf{x}\in\mathbb{R}^m\,\vert\,\mathbf{B}\,\mathbf{x}\leq\mathbf{b}\}$	$\hat{\mathbf{x}}=\mathrm{cen}(FSS)$	Analytical expression of $\hat{\mathbf{x}}$ not available
Observation errors	$\mathbf{e}$ is composed of stochastic error components ($\mathbf{v}$), and remaining systematic error components ($\mathbf{s}$).			
Observation uncertainty modeling[1]	Stochasticity: $E(\mathbf{v})=\mathbf{0}$, $D(\mathbf{v})=\Sigma=\mathbf{P}^{-1}$ Systematics: $[\mathbf{s}]=\mathbf{F}\cdot[\mathbf{d}]$	Stochasticity: $[v_i]=[-k_\alpha\sigma_i,k_\alpha\sigma_i]$, if $\mathbf{v}\sim\mathcal{N}(\mathbf{0},\mathbf{\Sigma})$ Systematics: $[\mathbf{s}]=\mathbf{F}\cdot[\mathbf{d}]$	Stochasticity: $[v_i]=[-k_\alpha\sigma_i,k_\alpha\sigma_i]$, if $\mathbf{v}\sim\mathcal{N}(\mathbf{0},\mathbf{\Sigma})$ Systematics: $[\mathbf{s}]=\mathbf{F}\cdot[\mathbf{d}]$	Stochasticity: $E(\mathbf{v})=\mathbf{0}$, $D(\mathbf{v})=\Sigma=\mathbf{P}^{-1}$ and $[v_i]=[-k_\alpha\sigma_i,k_\alpha\sigma_i]$, if $\mathbf{v}\sim\mathcal{N}(\mathbf{0},\mathbf{\Sigma})$ Systematics: $[\mathbf{s}]=\mathbf{F}\cdot[\mathbf{d}]$
Confidence region	Stochasticity: Confidence ellipse $\mathcal{E}_{1-\alpha}$ Systematics: Zonotope $\mathcal{Z}_D$ Combined: Extended confidence region $\mathcal{C}_E$	Confidence zonotope $\mathcal{Z}_s$[2]	FSS[3]	Constrained confidence region $\mathcal{C}_C$
Estimation error bound[1]	Stochasticity: $\|\varepsilon\|\leq\mathrm{rad}(\mathcal{E}_{1-\alpha})$ Combined: $\|\varepsilon\|\leq\mathrm{rad}(\mathcal{C}_E)$	$\|\varepsilon\|\leq\mathrm{diam}(\mathcal{P}_s)$	$\|\varepsilon\|\leq\mathrm{rad}(FSS)$	$\|\varepsilon\|\leq\max_{\mathbf{x}\in\mathcal{C}_\mathbf{C}}\|\hat{\mathbf{x}}-\mathbf{x}\|\leq\mathrm{diam}(\mathcal{C}_C)$

[1] The error bounds, especially with respect to stochasticity, may be associated with probabilities.
[2] A confidence zonotope is not a *statistically-viewed* confidence region but is defined as the nominal solution for the set estimator.
[3] The FSS is a key input to the set-based central estimator and set-constrained least-squares estimator. An example is the outcome of the set estimator. The determination of FSS may vary in complex systems, such as a "combination" of multiple set solutions, which is further discussed in Chap. 5.

- The least-squares estimator degenerates into a *mean* estimator:

$$\hat{\mathbf{x}}_{LS} \coloneqq (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{P}\,\mathbf{y} = \frac{\Sigma_{i=1}^{n} y_i}{n}. \tag{4.47}$$

- The set estimator degenerates into an interval estimator related to the *sample range*:

$$[\hat{\mathbf{x}}] = [\max(\mathbf{y}) - \Delta_s, \min(\mathbf{y}) + \Delta_s]. \tag{4.48}$$

- Taking the set solution $[\hat{\mathbf{x}}]$ from Eq. 4.48 as the input FSS, the set-based central estimator degenerates into a *midpoint* estimator, or termed *mid-range* estimator:

$$\hat{\mathbf{x}}_c = \frac{1}{2} \cdot (\max(\mathbf{y}) + \min(\mathbf{y}))\,, \tag{4.49}$$

  conditioning on $[\hat{\mathbf{x}}] \neq \emptyset$.

- The set-constrained least-squares (SCLS) estimator can be expressed with respect to the least-squares and set estimators:

$$\hat{\mathbf{x}}_{SC} = \begin{cases} \max(\mathbf{y}) - \Delta_s, \text{ i.e., lower bound of } [\hat{\mathbf{x}}], & \text{if } \hat{\mathbf{x}}_{LS} \leq \max(\mathbf{y}) - \Delta_s, \\ \hat{\mathbf{x}}_{LS}, & \text{if } \hat{\mathbf{x}}_{LS} \in [\hat{\mathbf{x}}], \\ \min(\mathbf{y}) + \Delta_s, \text{ i.e., upper bound of } [\hat{\mathbf{x}}], & \text{if } \hat{\mathbf{x}}_{LS} \geq \min(\mathbf{y}) + \Delta_s, \end{cases} \tag{4.50}$$

  conditioning on $[\hat{\mathbf{x}}] \neq \emptyset$.

Consequently, it becomes possible to evaluate and compare the estimators qualitatively:

- **Statistical property.** The classical least-squares estimator is optimal with the least variance of estimation and unbiasedness thanks to the best linear unbiased estimation principle. In contrast, the set-based central estimator can provide an unbiased, minimum-variance solution for uniformly distributed errors. It, as well as the set estimator, may be regarded as L estimators under certain conditions since they depend on the maximum and minimum values of the samples. The SCLS may be regarded as a compromise between the least-squares estimator and set-based central estimator.
- **Robustness.** None of the least-squares estimator, the set estimator, nor the set-based central estimator in the current problem is robust. They are sensitive to outliers. The set estimator and the set-based central estimator may return an empty solution in the presence of extreme outliers. In contrast, the classical least-squares estimator still outputs a point solution but loses its unbiasedness feature in the same situation. While the SCLS is desired to enhance the robustness of the classical least-squares estimator, it can benefit from the additional set constraint if and only if the interval error bounding is adequate (e.g., high confidence).

4.4.3 Statistical evaluation for a two-dimensional problem

This subsection uses a Monte Carlo approach to conduct a statistical evaluation, focusing on the two-dimensional positioning scenario introduced in Sec. 4.3.1. Various observation error models are employed to generate datasets (four measurements at each of 10^5 epochs).

Scenario definition. Five different scenarios are evaluated. Fig. 4.6 displays the simulated errors (black dotted curves) and established error bounds for different estimators in Scenarios 2-5 (blue-colored for interval bounds and red solid curves for Gaussian bounds), compared to the standard normal distribution (Scenario 1, red dashed curves). All the error bounds are adequate; however, they may be pessimistic for certain probabilities inherently.

- **Scenario 1:** Observation errors are identically and independently normally distributed, i.e., $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Interval bounds for the set estimator are determined at a confidence level $1-\alpha$, specifically $[-k_\alpha, k_\alpha]$ (cf. Eq. 4.27).
- **Scenario 2:** Observation errors follow a uniform distribution $\mathbf{e} \sim \mathcal{U}_{[-3,3]}$. Hence, all generated samples lie within the interval bounds $[-3, 3]$, and the VCM of the observations is computed as
$$\Sigma = \frac{(3-(-3))^2}{12} \cdot \mathbf{I} = 3 \cdot \mathbf{I}. \tag{4.51}$$
However, Eq. 4.51 is not adequate to establish an overbounding Gaussian distribution, as shown in Fig. 4.6(a). The standard deviation should be inflated to achieve CDF overbounding, cf. Sec. 2.4.2. The following evaluation adopts the overbounding sigma as $\sigma_{ob} = 2.3293$.
- **Scenario 3:** Observation errors follow identically and independently long-tailed distributions. A student's t-distributed error model $\mathbf{e} \sim t(\nu = 10)$ is used to represent this scenario. An overbounding sigma is obtained $\sigma_{ob} = 1.9307$. For the set estimator, the interval bounds are determined as $[-k_\alpha - b_{t-N}, k_\alpha + b_{t-N}]$, where the term b_{t-N} is defined for overbounding the tail distribution:
$$b_{t-N} = Q^{-1}_{t(\nu=10)}\left(\frac{1}{4\times 10^5}\right) - Q^{-1}\left(\frac{1}{4\times 10^5}\right) \approx 4.2474, \tag{4.52}$$
where $Q^{-1}(P)$ and $Q^{-1}_{t(\nu)}(P)$ denote the $(1-P)$ quantile of a normal distribution and a student's t-distribution of ν degree of freedom.
- **Scenario 4:** Observation errors are modeled as $\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2$. The component $\mathbf{e}_1$ follows a Gaussian distribution $\mathbf{e}_1 \sim \mathcal{N}(\mathbf{0}, 0.5^2 \cdot \mathbf{I})$, and $\mathbf{e}_2$ follows a uniform distribution $\mathbf{e}_2 \sim \mathcal{U}_{[-3,3]}$. Interval bounds for the set estimator are computed as $[-0.5 \cdot k_\alpha - 3, 0.5 \cdot k_\alpha + 3]$.
- **Scenario 5:** Observation errors simulate a hybrid model $\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2$. The component $\mathbf{e}_1$ follows a normal distribution $\mathbf{e}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The component $\mathbf{e}_2$ is a deterministic bias known to be upper bounded by an interval $[-3, 3]$. In the MC trials, this component is set to $e_2 = 3$ for generating the observation samples. The set estimator uses interval bounds $[-k_\alpha - 3, k_\alpha + 3]$.

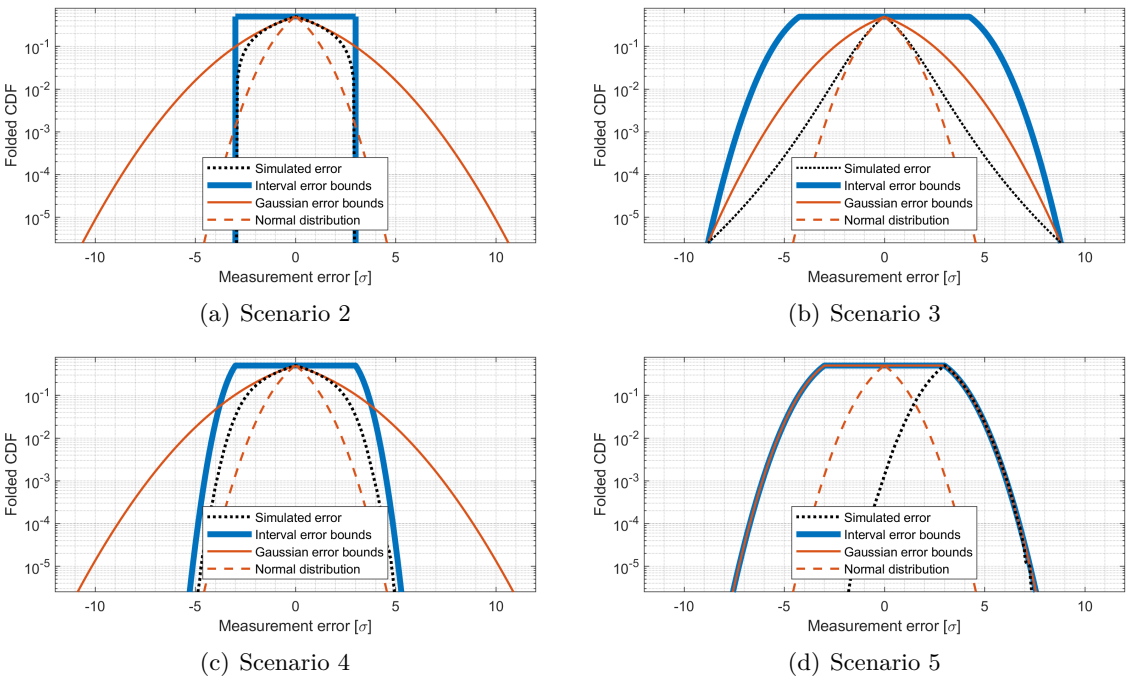
Furthermore, three scenarios are defined for the SCLS estimator based on the determination of set constraints, which utilizes observation interval bounds and the set estimator based on constraint satisfaction:

- **Scenario *a*:** The confidence requirement α is equally allocated to the least-squares and set estimators ($\alpha_p = \alpha_e$).
- **Scenario *b*:** More stringent confidence requirement for the set estimator ($\alpha_p < \alpha_e$).
- **Scenario *c*:** A deterministic set constraint is specified, i.e., the set estimator is not associated with any confidence requirement.

The confidence level is set to $1 - \alpha = 1 - 10^{-5}$.

The estimation error bounds are computed using the formulas in Tab. 4.4. Tab. 4.5 summarizes the interval bounds used for each estimator, with specific details for the SCLS estimator (e.g., the confidence requirement allocation). Particularly, interval bounds in Scenario *c* are determined by the maximum value of each observation in the MC trials.

Comparison of the Set, Central, and Least-Squares Estimators. The following performance metrics are focused on:



(a) Scenario 2 (b) Scenario 3

(c) Scenario 4 (d) Scenario 5

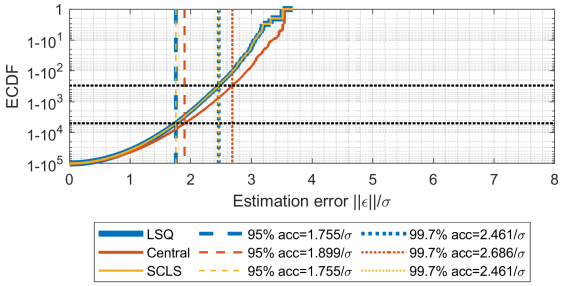
Figure 4.6: Scenarios of observation error modeling for statistical evaluation in the form of folded CDFs. Four subfigures show Scenarios 2-5, where the interval error bounds and Gaussian error bounds are compared with the simulated errors and standard normal distribution (i.e., Scenario 1). For illustration in subfigures (b) - (d), interval error bounds are shown as a combination of two Gaussian distributions: $\mathcal{N}(-\Delta, \sigma_{ub}^2)$ (negative side of the folded CDF) and $\mathcal{N}(\Delta, \sigma_{ub}^2)$ (positive side of the folded CDF), with $[-\Delta, \Delta]$ for the interval-bounded error component and σ_{ub} for the stochastic error component.

- 95% and 99.7% percentiles of estimation errors, which measure **accuracy**. Three point estimators are evaluated. Smaller values represent better accuracy.
- Maximum value of the **estimation error bounds**, comparing four estimators (note: the least-squares estimator has a fixed estimation error bound that depends on the required confidence level). An estimator with smaller error bounds is preferred.

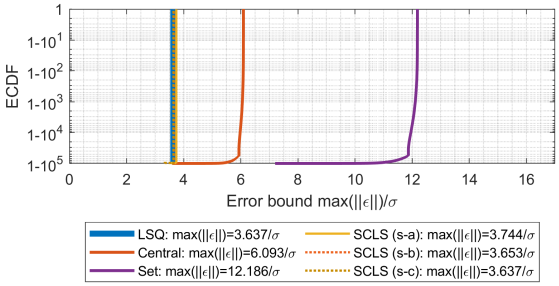
These metrics are assessed based on the ECDFs for estimation errors and their error bounds in Fig. 4.7.

Tab. 4.6 ranks the performance of each estimator across the evaluated scenarios. The key insights are as follows:

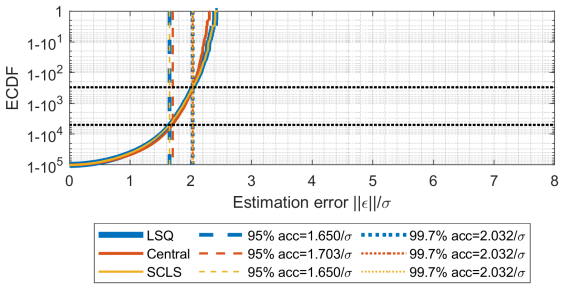
- In an ideal situation, such as Scenario 1, where the systematic effects are negligible and the stochastic errors are perfectly modeled (i.e., Gaussian), pure stochastic handling shows superior performance in both accuracy and error bounds. This supports the conclusion that pure stochastic handling through Gaussian overbounding and least-squares estimator is optimal under such conditions.
- When systematics contribute to a long-tailed error distribution (Scenario 3), stochastic approaches continue to outperform other approaches. Signal-In-Space Range Error (SISRE) exhibits such behaviors, as shown in Fig. 2.6(a).
- For errors yielding interval bounds, deterministic approaches become advantageous, e.g., Scenario 2. The set-based central estimator improves the 99.7% accuracy and error bounding compared to the least-squares estimator.



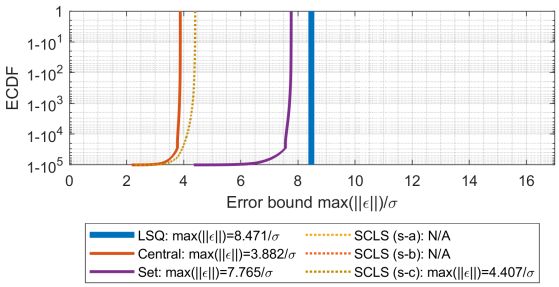
(a) Accuracy comparison for Scenario 1



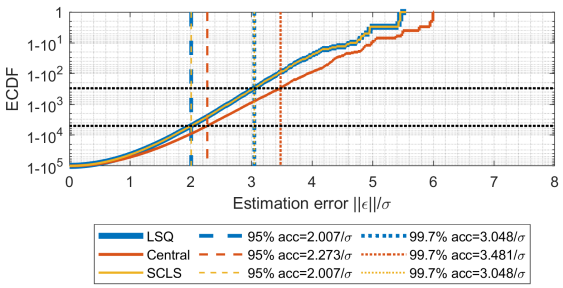
(b) Error bound comparison for Scenario 1



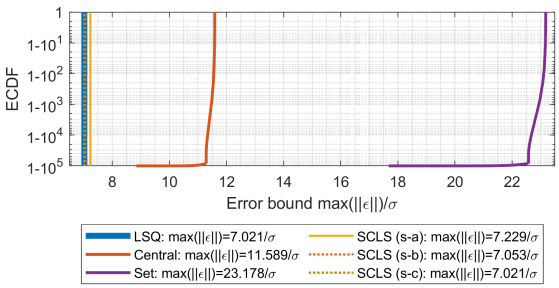
(c) Accuracy comparison for Scenario 2



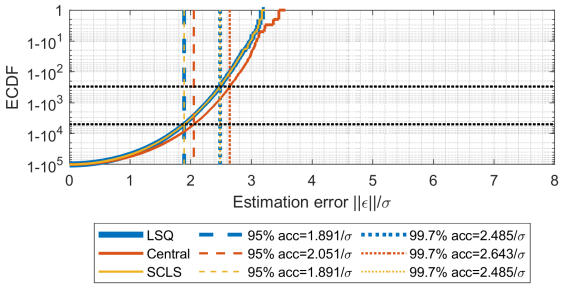
(d) Error bound comparison for Scenario 2



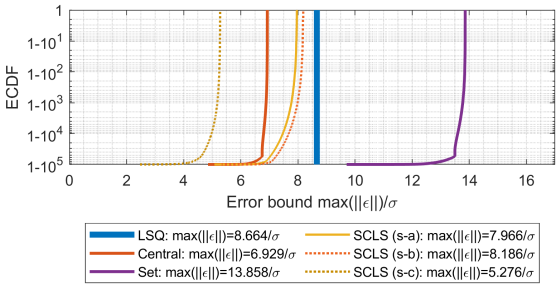
(e) Accuracy comparison for Scenario 3



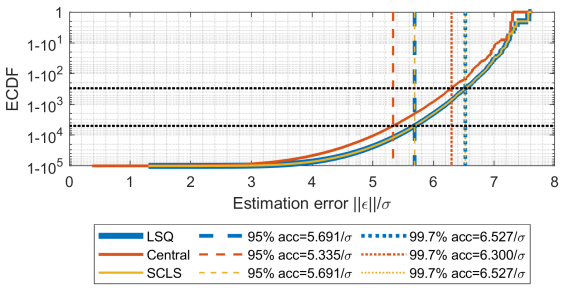
(f) Error bound comparison for Scenario 3



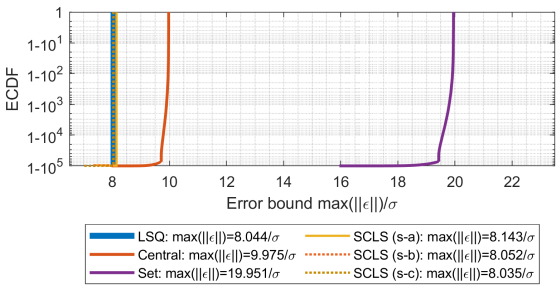
(g) Accuracy comparison for Scenario 4



(h) Error bound comparison for Scenario 4



(i) Accuracy comparison for Scenario 5



(j) Error bound comparison for Scenario 5

Figure 4.7: Statistical evaluation for different estimators across five scenarios. 10^5 Monte Carlo trials are conducted for each scenario. The ECDFs for estimation errors and error bounds are shown.

Table 4.5: Interval bounds used in the statistical evaluation for different estimators.

Estimator		**Interval bounds**				
		Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Least-squares		N/A[1]	N/A[1]	N/A[1]	N/A[1]	3.0
Set & set-based central		4.7081	3.0	8.9555	5.3541	7.7081
SCLS	Scenario a ($\alpha_p = 5 \cdot 10^{-6}$)	4.8475	N/A[2]	9.0949	5.4238	Set: 7.8475 LSQ: 3.0
	Scenario b ($\alpha_p = 1 \cdot 10^{-6}$)	5.1577	N/A[2]	9.4051	5.5789	Set: 8.1577 LSQ: 3.0
	Scenario c[3]	$\begin{bmatrix}4.7\\4.3\\4.3\\4.2\end{bmatrix}$	$\begin{bmatrix}3.0\\3.0\\3.0\\3.0\end{bmatrix}$	$\begin{bmatrix}8.6\\7.8\\9.7\\9.4\end{bmatrix}$	$\begin{bmatrix}3.5\\3.6\\3.5\\3.7\end{bmatrix}$	Set: $\begin{bmatrix}7.3\\7.3\\7.3\\7.1\end{bmatrix}$ LSQ: 3.0

[1] Interval error bounding is not applied and, hence, "N/A" is indicated.
[2] Interval bounds are adequate in the case of uniform distributions and, hence, not necessarily associated with probabilities, making these two scenarios invalid.
[3] The interval bounds are determined by the maximum value of the samples.

- Scenario 4 represents a more complex and realistic situation with stochastic errors plus remaining systematic errors that dominate the total budget. The overbounding Gaussian distribution tends to pessimistic overbounding at tail probabilities due to excessive kurtosis, as shown in Fig. Fig. 4.6(c). In contrast, the interval error bounds provide tighter overbounding, and subsequently, the central estimator demonstrates smaller estimation error bounds than the least-squares estimator.
- When biases are only upper-bounded by intervals, interval bounds can still be integrated into the least-squares estimator to enhance estimation error bounding (cf. Sec. 4.2.1). Interstingly, in Scenario 5, the central estimator outperforms the least-squares estimator in terms of accuracy, while the least-squares estimator retains an advantage in error bounding. This illustrates the complementary nature of the two approaches in handling hybrid error models.
- The set estimator shows consistent error bounding performance relative to the central estimator, i.e., it offers error bounds that are double those of the central estimator. Hence, suggestions on applying the central estimator drawn from the evaluation hold for the set estimator as well.

Analysis for SCLS. In this evaluation, the accuracy performance of SCLS remains unchanged across three scenarios (*a, b*, and *c*) for each observation error model. This is due to two factors: (i) the simulation assumes fault-free, where the observation errors are adequately bounded, and (ii) a fairly strict confidence level for estimation. Consequently, the set constraint is not active in determining point solutions. The corresponding ECDFs are displayed in Fig. 4.7 (a), (c), (e), (g), and (i) for Scenarios 1-5.

In contrast, the estimation error bounds vary depending on the scenarios, influenced by both the observation error bounding (Scenarios 1-5) and confidence requirement allocation (Scenarios *a*-*c*). Key observations include:

Table 4.6: Performance comparison for different estimators in terms of estimation accuracy and error bounds. The number indicates the performance ranking under various scenarios. The best-performing estimator in each case is highlighted in bold. SCLS under Scenario *c* is involved in the comparison.

Scenario	Observation errors	Performance metric	Estimator			
			LSQ	*set*	*central*	*SCLS*
1	$\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$	accuracy (95%)	**1**	/	3	**1**
		accuracy (99.7%)	**1**	/	3	**1**
		error bounds*	**1**	4	3	2
2	$\mathbf{e} \sim \mathcal{U}_{[-3,3]}$	accuracy (95%)	**1**	/	3	**1**
		accuracy (99.7%)	2	/	**1**	2
		error bounds*	4	3	**1**	2
3	$\mathbf{e} \sim \mathbf{t}(\nu = 10)$	accuracy (95%)	**1**	/	3	**1**
		accuracy (99.7%)	**1**	/	3	**1**
		error bounds*	**1**	4	3	**1**
4	$\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2$ with	accuracy (95%)	**1**	/	3	**1**
	$\mathbf{e}_1 \sim \mathcal{N}(\mathbf{0}, 0.5^2 \times \mathbf{I})$	accuracy (99.7%)	**1**	/	3	**1**
	$\mathbf{e}_2 \sim \mathcal{U}_{[-3,3]}$	error bounds*	**3**	4	2	**1**
5	$\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2$ with	accuracy (95%)	2	/	**1**	2
	$\mathbf{e}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$	accuracy (99.7%)	2	/	**1**	2
	$\mathbf{e}_2 \in [-3, 3]$	error bounds*	2	4	3	**1**

* The estimation error bounds for set, central, and SCLS estimators are compared with respect to their maximum value.

- SCLS consistently benefits from both central and least-squares estimators, showing moderate performance in terms of accuracy and error bounds for all scenarios. It balances the strengths of both approaches without significantly sacrificing any performance metric.
- The confidence requirement allocation for SCLS directly impacts the interval bounds used for the set constraint, which in turn affects the SCLS error bounds. For instance, in Scenario 4 with dominant systematic errors, Scenario *b* (with the strictest confidence requirement for the set constraint) results in the largest interval and, subsequently, the largest estimation error bound. In Scenarios *a* and *c*, smaller interval bounds result in correspondingly reduced error bounds, visible in the ECDFs of estimation error bounds in Fig. 4.7(h). In other cases, the set constraint may be inactive in error bounding, and, as a result, the SCLS error bounds are solely dependent on the confidence requirement allocated to the least-squares estimator.
- Deterministic interval bounds are recommended to be used in SCLS when available. In Scenario *c*, interval bounds are derived from the maximum value of observation samples, leading to smaller estimation error bounds compared to the other scenarios (*a* and *b*). This finding is independent of observation error models. Particularly, Scenario 4 showcases the most notable improvement, as the other scenarios tend to overbound the observation errors conservatively.

Concluding Remarks. Four estimators – the classical least-squares estimator, set estimator, set-based central estimator, and set-constrained least-squares (SCLS) – were evaluated and compared for a two-dimensional example. Since each estimator operates under differ-

ent assumptions regarding observation errors, their performance was analyzed across various scenarios using metrics such as accuracy and error bounding.

The least-squares estimator excels in handling purely stochastic errors. In the case of pure Gaussian-distributed errors and long-tailed distributed errors, using the least-squares estimator with Gaussian overbounding led to better accuracy and error bounding compared to the other estimators.

The set estimator based on constraint satisfaction provides a set solution using interval error bounds in the observation domain rather than conventional point solutions. Its accuracy was not evaluated, and the estimation error bounds, accounting for all possible set elements, were not as small as those of the other estimators.

The set-based central estimator is designed to offer the minimum estimation error bounds providing observational interval bounds. When using the outcome of the set estimator as input (i.e., the feasible solution set (FSS)), it outperformed the classical least-squares estimator in both accuracy and error bounding for uniformly distributed errors (i.e., with known interval bounds) and hybrid errors dominated by interval-bounded components (e.g., significant remaining systematic errors).

Furthermore, the proposed SCLS estimator, combining the strengths of the least-squares principle with set constraints, demonstrated robust performance. Its effectiveness depends on the observation error model and the confidence requirement allocation. When both are carefully balanced, SCLS offers a flexible and reliable solution, as evidenced in various scenarios.

In conclusion, each estimator has strengths tailored to specific error models, and the choice of estimator should be aligned with the nature of the observation errors and the trade-off between accuracy and error bounding.

5

Interval-based GNSS receiver autonomous integrity monitoring

Classical GNSS integrity monitoring methods, including the (Advanced) RAIM, use residual-based (RB) or solution separation (SS) methods for fault detection and exclusion (FDE). While being widely implemented in aviation (RTCA/DO-229F, 2020; Joerger et al., 2012; Blanch et al., 2019a), integrity concepts have not yet achieved a similar level of maturity for land navigation. Recent years have seen investigations into detectors and estimators utilizing different strategies to fill this gap, e.g., Blanch and Walter (2021) and Wendel (2022).

Interval or set-theoretic methods, representing uncertainty as intervals or sets, explore the linear uncertainty propagation in localization problems, computing confidence zones in which the user is claimed to be located with a given confidence or risk (Moore et al., 2009; Meizel et al., 2002; Drevelle and Bonnifait, 2009). This chapter, partially stemming from prior works (Su and Schön, 2022a; Su et al., 2023), aims to evaluate the interval or set-theoretic methods within the integrity context and to develop a practical approach for autonomous integrity monitoring. The remainder of this chapter first formulates the set-based detector, evaluates its performance in a probabilistic context for benchmark examples, and compares it to classical test theory-based methods. The second section focuses on the proposed interval-based approach utilizing set-based FDE. The architecture is developed based on the evaluation of the probability of loss of integrity (LOI), design of optimal estimator, and solutions to Protection Level (PL) computation. Its performance is then evaluated and compared to conventional approaches through analytical analysis and Monte Carlo simulation.

5.1 Set-based fault detection

5.1.1 Introduction

Fault detection (FD) and fault exclusion (FE) are critical for ensuring the reliability and safety of complex systems across various domains, such as aerospace, process controls, automotive and manufacturing, etc. (Katipamula and Brambley, 2005). In GNSS applications, a faulty measurement or a satellite failure that may harm the navigation solution must be detected, and subsequently, a procedure may be activated to identify and exclude the fault to maintain a reliable navigation solution (Wang and Ober, 2009). Conventional techniques often rely on stochastical approaches. For instance, the detection, identification and adaptation (DIA) theory is widely adopted for data snooping in geodetic applications (Teunissen, 2018), and the RB and SS detectors are utilized in the RAIM and ARAIM. Details on these two detectors are provided in Sec. 2.4.3.

Efforts concerning FDE for navigation integrity purposes over the past decades have focused on two main methodological aspects:

- **Reducing computation load** by addressing certain issues related to FDE procedures;
- **Exploiting detector characteristics** to precisely model or tightly bound associated probabilities.

For example, reducing the number of monitored fault modes can boost the computation efficiency of SS approach without substantially degrading the navigation performance (Walter et al., 2014; Ge et al., 2017; Blanch et al., 2018; Wang et al., 2023). By formulating FDE as an optimization problem of an underdetermined linear system, Yang and Sun (2020) applied the least absolute shrinkage and selection operator and solved the problem relatively efficiently. Zhang and Papadimitratos (2021) proposed to accelerate the FE function by focusing on specific subset pairs of interest. Moreover, greedy search-based approaches have been explored to reduce the computational complexity when dealing with a large number of faults, applicable to both GNSS standalone systems (Blanch et al., 2015b; Wendel, 2022) and multi-sensor fusion scenarios (Zhu et al., 2023).

Concerning the second aspect, various studies aim at accurately determining the worst-case fault that results in the largest possible integrity risk (Milner and Ochieng, 2011; Jiang and Wang, 2014). Other efforts include refining the upper bounds for position error distributions (Liu et al., 2022; Racelis and Joerger, 2022), and precisely evaluating the risks or optimizing risk allocation (Blanch et al., 2010, 2013; Joerger and Pervan, 2016; Zhai et al., 2018).

Despite these advancements, traditional strategies predominantly rely on stochastic modeling of measurement errors, using adapted or combined RB and SS detectors. Recently, the set-based method, or named set-membership approach, has emerged as a promising alternative for FD. This approach explicitly computes the set of parameters or states consistent with the measurements; a fault is declared to have occurred once a measurement is found to be inconsistent with this set (Puig et al., 2007).

A number of studies have demonstrated the effectiveness of set-based detection across different fields. For example, approximating the parameter uncertainty by different classes of set such as zonotopes (Ingimundarson et al., 2009; Samada et al., 2023), polytopes (Blesa et al., 2012), ellipsoids (Watkins and Yurkovich, 1996) and parallelotopes (Kesavan and Lee, 2001), set-based detection has been successfully implemented in wind turbine monitoring (Blesa et al., 2011), sewer network monitoring in Barcelona (Puig and Blesa, 2013), and aerospace applications (Pons et al., 2008). In the context of satellite navigation, Drevelle and Bonnifait (2009) demonstrated the potential of set-based methods to enhance positioning performance. Despite these developments, a rigorous formulation of set-based detection tailored to strict integrity requirements remains an open question.

The remainder of this section presents a formal formulation of set-based detection for GNSS positioning. This is followed by analytical evaluations within a probabilistic framework using benchmark examples, and a Monte Carlo assessment of two-dimensional scenarios. The performance of the set-based method is analyzed and compared against classical detection techniques, leading to a proposed enhancement through a novel weighting strategy for improved robustness.

5.1.2 Basics of set-emptiness method

The detection problem for GNSS navigation can be modeled as follows:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e} + \mathbf{f}, \tag{5.1}$$

where $\mathbf{y}$, $\mathbf{A}$, $\mathbf{x}$ and $\mathbf{e}$ are introduced in Eq. 4.2. The additional $n \times 1$ vector of unknown faults $\mathbf{f}$ represents the unmodeled biases to be detected. $\mathbf{f} = \mathbf{0}$ if the measurements are fault-free.

To construct a detector, a vector of thresholds $\pm\boldsymbol{\Delta}_d$ or intervals $[-\boldsymbol{\Delta}_d, \boldsymbol{\Delta}_d]$ is defined for an acceptable amount of observation errors. The primary objective is to identify the following conditions:

$$|\mathbf{e}| \leq \boldsymbol{\Delta}_d, \text{ while} \tag{5.2}$$

$$|e_j + f_j| > \Delta_{d,j} \text{ if } f_j \neq 0, \text{ with } e_j \in \mathbf{e}, f_j \in \mathbf{f}, \Delta_{d,j} \in \boldsymbol{\Delta}_d. \tag{5.3}$$

This is a straightforward way of detecting faulty measurements using intervals, aligning with the *direct test* introduced in Puig (2010). However, The errors $\mathbf{e}$ are not accessible at the measurement level in GNSS problems. To address this, Eq. 5.2 is reformulated as

$$\begin{aligned} -\boldsymbol{\Delta}_d \leq \mathbf{e} \leq \boldsymbol{\Delta}_d \\ \Rightarrow \mathbf{A}\mathbf{x} - \boldsymbol{\Delta}_d \leq \mathbf{A}\mathbf{x} + \mathbf{e} \leq \mathbf{A}\mathbf{x} + \boldsymbol{\Delta}_d, \end{aligned} \tag{5.4}$$

and Eq. 5.3 is rewritten as

$$\begin{aligned} e_j + f_j > \Delta_{d,j} \text{ or } e_j + f_j < -\Delta_{d,j} \\ \Rightarrow \mathbf{A}^j\mathbf{x} + e_j + f_j > \mathbf{A}^j\mathbf{x} + \Delta_{d,j} \text{ or } \mathbf{A}^j\mathbf{x} + e_j + f_j < \mathbf{A}^j\mathbf{x} - \Delta_{d,j} \\ \text{if } f_j \neq 0, \text{ with } e_j \in \mathbf{e}, f_j \in \mathbf{f}, \Delta_{d,j} \in \boldsymbol{\Delta}_d, \end{aligned} \tag{5.5}$$

where $\mathbf{A}^j$ denotes the j-th row of matrix $\mathbf{A}$.

The desired detector should differentiate between fault-free and faulty conditions:

- **Fault-free case** ($\mathbf{f} = \mathbf{0}$). Substituting $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$ in Eq. 5.4 leads to the following inequalities:

$$\mathbf{A}\mathbf{x} - \boldsymbol{\Delta}_d \leq \mathbf{y} \leq \mathbf{A}\mathbf{x} + \boldsymbol{\Delta}_d \Rightarrow \begin{cases} \mathbf{A}\mathbf{x} \leq \mathbf{y} + \boldsymbol{\Delta}_d \\ -\mathbf{A}\mathbf{x} \leq -\mathbf{y} + \boldsymbol{\Delta}_d \end{cases}. \tag{5.6}$$

- **Faulty cases** ($\mathbf{f} \neq \mathbf{0}$). The inequalities in Eq. 5.5 are reformulated as

$$\begin{aligned} \mathbf{y} \not\leq \mathbf{A}\mathbf{x} + \boldsymbol{\Delta}_d \text{ or } \mathbf{y} \not\geq \mathbf{A}\mathbf{x} - \boldsymbol{\Delta}_d \\ \Rightarrow -\mathbf{A}\mathbf{x} \not\leq -\mathbf{y} + \boldsymbol{\Delta}_d \text{ or } \mathbf{A}\mathbf{x} \not\leq \mathbf{y} + \boldsymbol{\Delta}_d, \end{aligned} \tag{5.7}$$

  where $\not\leq$ and $\not\geq$ denote the violation of component-wise vector inequalities, cf. the corresponding entry in the *Notations and Nomenclatures* section for details.

These inequalities represent the expected binary outcomes of fault detection, essentially checking the validity of Eq. 5.6. Notably, the detection operates in the state domain ($\mathbf{x} \in \mathbb{R}^m$).

Solving the constraint satisfaction problem in Eq. 5.6 results in a polytope set solution:

$$\mathcal{P}_d(\mathbf{B}, \mathbf{b}) = \left\{ \mathbf{x} \in \mathbb{R}^m \;\middle|\; \mathbf{B}\mathbf{x} \leq \mathbf{b}, \mathbf{B} = \begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \end{bmatrix} \in \mathbb{R}^{2n \times m}, \mathbf{b} = \begin{bmatrix} \mathbf{y} + \boldsymbol{\Delta}_d \\ -\mathbf{y} + \boldsymbol{\Delta}_d \end{bmatrix} \in \mathbb{R}^{2n} \right\}. \tag{5.8}$$

Detection occurs when the detection polytope is empty ($\mathcal{P}_d = \emptyset$), signifying that no feasible solution to the constraint satisfaction problem is available. $\mathcal{P}_d$ can be interpreted as an "inconsistency area" defined by specific interval constraints and satellite geometry.

Geometrically, each interval forms a "slab" ($\mathcal{S}_j$), perpendicular to the j-th LOS:

$$\mathcal{S}_j = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{A}^j\mathbf{x} \in [y_j - \Delta_{d,j}, y_j + \Delta_{d,j}]\}. \tag{5.9}$$

It follows that $\mathcal{P}_d$ is the intersection of all these "slabs".

Similar to Eq. 4.19, a confidence zonotope can be obtained in an ideal situation of error-free ($\mathbf{e} = \mathbf{0}$) and fault-free ($\mathbf{f} = \mathbf{0}$), regarded as a nominal solution:

$$\begin{aligned} \mathcal{Z}_d &= \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{A}\mathbf{x} \leq \mathbf{y}^* + \boldsymbol{\Delta}_d, -\mathbf{A}\mathbf{x} \leq -\mathbf{y}^* + \boldsymbol{\Delta}_d\} \\ &= \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{A}(\mathbf{x} - \mathbf{x}^*) \leq \boldsymbol{\Delta}_d, -\mathbf{A}(\mathbf{x} - \mathbf{x}^*) \leq \boldsymbol{\Delta}_d\}, \end{aligned} \tag{5.10}$$

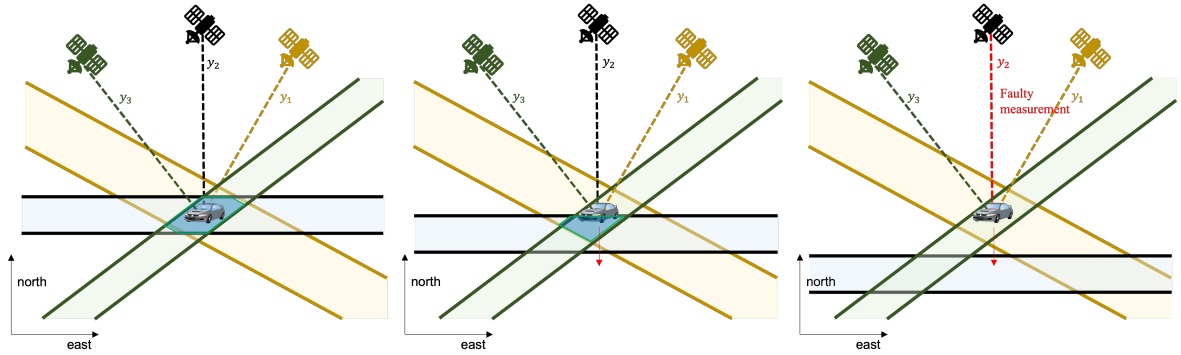
which is centered at the unknown true position $\mathbf{x}^*$.

The geometrical difference of $\mathcal{P}_d$ and $\mathcal{Z}_d$ indicates the impact of observational consistency in the position domain and may be quantified via an "inconsistency measure" based on their relative volume Vol_r (Dbouk and Schön, 2020):

$$Vol_r = \frac{Vol_Z - Vol_P}{Vol_Z}, \tag{5.11}$$

where Vol_Z and Vol_P denote the volume (area) of the zonotope and polytope, respectively. A lower inconsistency measure is expected for higher reliability of the solution.

Fig. 5.1 conceptualizes the two-dimensional slabs and detection polytopes in various scenarios:



(a) Nominal solution set as zonotope (b) Non-empty polytope (c) Empty polytope

Figure 5.1: Conceptual sketch of the detection polytope (blue), represented as the intersection of two-dimensional slabs.

- Fig. 5.1(a) depicts the nominal zonotope solution $\mathcal{Z}_d$;
- In reality, $\mathbf{e} \neq \mathbf{0}$ as in Figure 5.1(b), where $\mathcal{P}_d$ is deformed and shifted from the nominal situation to the indicated direction, dependent on both geometry and measurement errors.
- An unknown fault occurs in Fig. 5.1(c) and results in no intersection (detection triggered), i.e., no feasible point solution for the inequalities in Eq. 5.6.

This thesis takes *set emptiness* as the test criterion, alleviating the need for scalar test thresholds typical of classical testing theory-based approaches.

5.1.3 Performance evaluation: Benchmark problem

Introduction

The set-based detector is designed to address the fundamental question:

Are the measurements considered to be self-consistent, given the interval-based uncertainty modeling?

Evaluating the detection performance entails answering two key questions:

1. How likely is it that the detection is mistakenly triggered by properly modeled errors?
2. What is the failure rate of detection in the presence of unmodeled or unbounded faults?

Statistical hypothesis testing provides a framework for checking the statistical significance of information about parameters (Koch, 1999). For our detection problem of interest, the following hypotheses are defined:

- **Null hypothesis (H_0).** All observation error components are exhaustively modeled, and the resulting interval bounds are adequate.
- **Alternative hypothesis (H_A).** In addition to interval-bounded error components, unmodeled or unbounded faults exist and may bias the state estimator.

The task of detection, therefore, becomes deciding whether to reject the null hypothesis in favor of the alternative. This decision-making process involves assessing the risk of two types of errors (Teunissen, 2006):

- ***Type I* error** occurs if H_0 is rejected erroneously and it is actually true, alternatively named *false alert (FA)*.
- ***Type II* error** occurs due to the failure of rejecting H_0 when it is false, also known as *missed detection (MD)*.

Minimizing *Type I* errors ensures that FA is rare, meaning the detector does not overreact to "normal" variations due to properly modeled errors. Conversely, reducing *Type II* errors enhances the detector's ability to identify actual faults, thus decreasing the chances of MD.

However, minimizing both error types simultaneously is impossible due to their inherent trade-off. Furthermore, in the context of GNSS integrity, both the magnitude and distribution of potential faults are typically unknown, complicating the task of assessing the impact of FD on navigation continuity and integrity.

This section aims to quantitatively evaluate the performance of the set-based detector in terms of the FA and MD probabilities. For clarity and simplicity, a scalar benchmark example is analyzed, assuming a standard normal distribution for the error components. The performance of the set-based approach is compared with traditional methods, including RB and SS approaches.

Benchmark problem

A simplified scenario is considered to demonstrate the detection principle and performance in a probabilistic context:

1. Scalar state estimation problem;
2. Exclusive Gaussian-distributed errors in the measurement model.

A benchmark example is used, cf. Joerger et al. (2013), without making unnecessarily complicated context. The problem is configured with three measurements:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e} + \mathbf{f},\ \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T. \tag{5.12}$$

The following assumptions are made:

- $\mathbf{y}^* = \mathbf{A}\,\mathbf{x}^* = \mathbf{0}$: the initial guess of the state is overlapped with the truth;
- $\mathbf{e} \sim N(\mathbf{0}, \mathbf{I})$: the errors follow standard normal distributions;
- Equal magnitude of the detector intervals $[-\boldsymbol{\Delta}_d, \boldsymbol{\Delta}_d]$;
- The fault vector $\mathbf{f}$ represents three single-measurement faults, corresponding to three alternative hypotheses H_i, with unknown fault magnitude f_i:

$$\mathbf{f} = \begin{bmatrix} f_1 \\ 0 \\ 0 \end{bmatrix} \text{ or } \mathbf{f} = \begin{bmatrix} 0 \\ f_2 \\ 0 \end{bmatrix} \text{ or } \mathbf{f} = \begin{bmatrix} 0 \\ 0 \\ f_3 \end{bmatrix}. \tag{5.13}$$

Subsequently, the following inequality system is obtained for the set-based detection problem in Eq. 5.6:

$$\begin{cases} x \le y_1 + \Delta_d, & -x \le -y_1 + \Delta_d, \\ x \le y_2 + \Delta_d, & -x \le -y_2 + \Delta_d, \\ x \le y_3 + \Delta_d, & -x \le -y_3 + \Delta_d, \end{cases} \tag{5.14}$$

which results in a detection polytope, i.e., an interval:

$$\mathcal{P}_d := [\underline{x}, \overline{x}] = [\max(y_1, y_2, y_3) - \Delta_d, \min(y_1, y_2, y_3) + \Delta_d]. \tag{5.15}$$

To retain its non-emptiness, the following inequality must be ensured:

$$\max(y_1, y_2, y_3) - \min(y_1, y_2, y_3) \le 2\Delta_d. \tag{5.16}$$

Detection is triggered once the difference value exceeds $2\Delta_d$.

Test statistic formulation

Denote the difference value in Eq. 5.16 by W and extend the question to n measurements $\mathbf{y} = [y_1, ..., y_n]^T$. Then, W and its criterion can be expressed as follows:

$$W := \max(y_1, ..., y_n) - \min(y_1, ..., y_n) \le 2\Delta_d, \tag{5.17}$$

which can be interpreted as an analog "test statistic". By nature, this formulation represents an application of *order statistics* (David and Nagaraja, 2003):

> *Let $Y_1, Y_2, ..., Y_n$ be independent random variables drawn from a population with CDF $F_N(y)$ and PDF $f_N(y)$. Then the CDF of $Y_{(r:n)}$ (r-th order statistic, $r = 1, ..., n$) is given by*
>
> $$\begin{aligned} F_{Y_{(r:n)}}(y) =& P\left(Y_{(r:n)} \le y\right) \\ =& P\left(\textit{at least } r \textit{ of } Y_1, Y_2, ..., Y_n \textit{ are } \le y\right) \\ =& \sum_{i=r}^{n} \binom{n}{i} \cdot \left(F_N(y)\right)^i \cdot \left(1 - F_N(y)\right)^{n-i}. \end{aligned} \tag{5.18}$$
>
> *The binomial probability in the summand represents the probability that exactly r of the Y_i's are no greater than y.*

Therefore, the test statistic W is the difference of the n-th order statistic (maximum, $Y_{(n:n)}$) and the first order statistic (minimum, $Y_{(1:n)}$), i.e., the *sample range*:

$$W = Y_{(n:n)} - Y_{(1:n)}. \tag{5.19}$$

Its PDF and CDF, under H_0, are

$$\begin{aligned} f_W(w) =& n(n-1) \int_{-\infty}^{\infty} f_N(y) \cdot \Big(F_N(y+w) - F_N(y)\Big)^{n-2} \cdot f_N(y+w)\,\mathrm{d}y, \\ F_W(w) =& n \int_{-\infty}^{\infty} f_N(y) \cdot \Big(F_N(y+w) - F_N(y)\Big)^{n-1}\,\mathrm{d}y. \end{aligned} \tag{5.20}$$

Fig. 5.2 presents the PDF and CDF of W with various Gaussian-distributed samples n. The derivation of the above equations is given in Annex. A.2.2, summarized based on textbooks such as David and Nagaraja (2003).

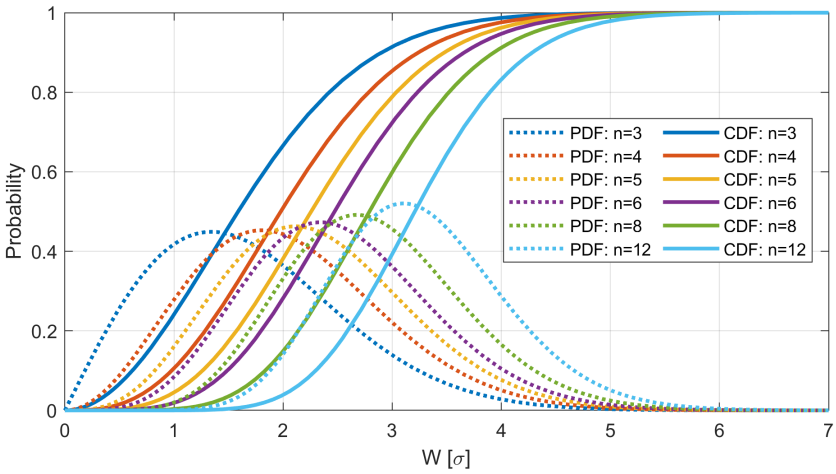


Figure 5.2: PDF and CDF of test statistic $W := Y_{(n:n)} - Y_{(1:n)}$ with various Gaussian-distributed samples n.

False alert probability

The probability of an empty intersection (a detection event) under H_0, i.e., false alert (FA), can be determined by

$$\begin{aligned} P(\mathrm{FA}) :=& P(\text{empty intersection} \Rightarrow W \geq 2\Delta_d \mid H_0) \\ =& \int_{2\,\Delta_d}^{\infty} f_W(w)\mathrm{d}w \\ =& 1 - F_W(2\Delta_d). \end{aligned} \tag{5.21}$$

To validate Eq. 5.21, Monte Carlo simulations are carried out with 10^7 trials for the benchmark problem. The FA occurrence is counted as Δ_d varies within the range of $[1.5\,\sigma,\, 4\,\sigma]$, compared to analytically determined probabilities using Eq. 5.21. In Fig. 5.3, the two curves, representing the results of the two approaches, are overlapped, verifying the evaluation.

Alternatively, if $P(\mathrm{FA})$ is limited by a required value, solving Eq. 5.21 results in a minimal Δ_d. For instance, $\Delta_d = 3.608\,\sigma$ for $P(\mathrm{FA}) = 1.003 \times 10^{-6}$, as indicated in Fig. 5.3. Fig. 5.4(a) showcases the relations between Δ_d/σ, n (the number of measurements) and $P(\mathrm{FA})$:

- Involving more measurements increases $P(\mathrm{FA})$ if Δ_d/σ is unchanged;
- Given a fixed number of measurements, using wider intervals tends to reduce $P(\mathrm{FA})$.

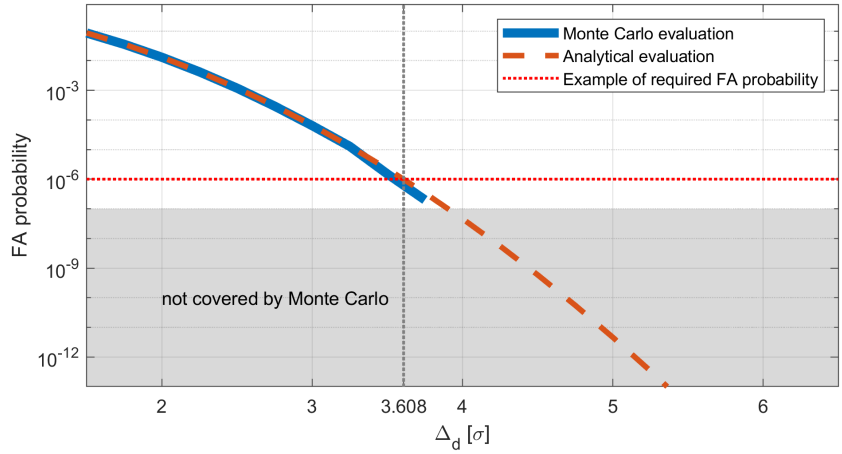
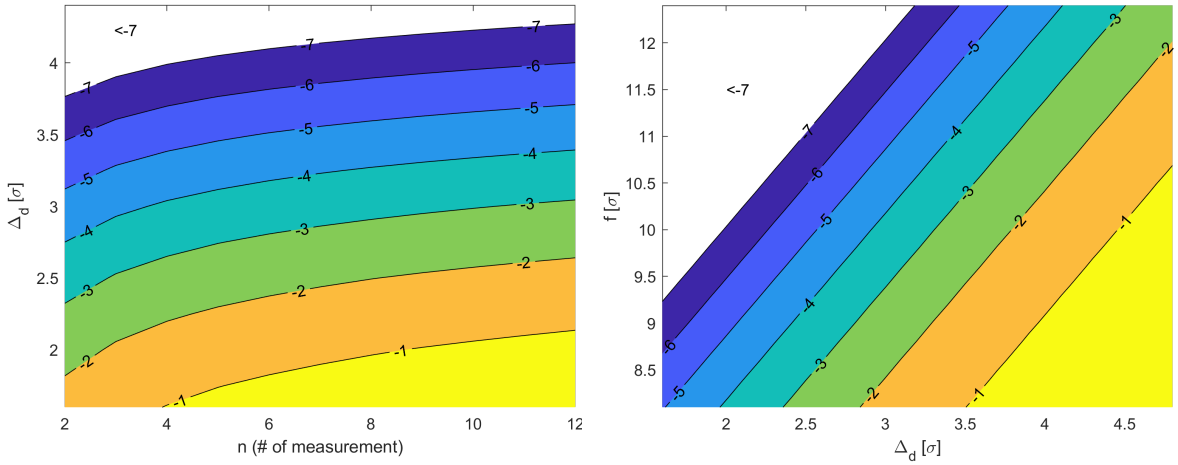


Figure 5.3: FA probabilities determined by analytical and Monte Carlo methods for the benchmark example.



(a) FA probabilities: $\log_{10}(P(\text{FA}))$

(b) MD probabilities: $\log_{10}(P(\text{ND} \mid \mathbf{f}, n = 3))$

Figure 5.4: Analytically determined FA and MD probabilities for the simplified scalar-state problems as contour plots.

Missed detection probability

Under an alternative hypothesis H_i, where the i-th measurement is faulty with the fault magnitude of f_i, a MD event occurs when the intersection remains non-empty. The MD probability is given by

$$\begin{aligned} P(\text{MD}) :=& P(\text{non-empty intersection} \Rightarrow W < 2\,\Delta_d \mid H_i) \\ =& \int_0^{2\,\Delta_d} f_W(w)\mathrm{d}w \\ =& F_W(2\,\Delta_d). \end{aligned} \tag{5.22}$$

In this case, the detection criterion remains, but the measurements no longer follow *independent identically distributed (iid)* Gaussian distributions. Instead, the i-th measurement may be considered to be drawn from a population with a CDF $F_N(y, f_i)$ and PDF $f_N(y, f_i)$ (i.e., a Gaussian distribution with non-zero mean $f_i \neq 0$). Consequently, the CDF $F_W(w)$ and PDF $f_W(w)$ of the test statistic W under H_i are influenced by the fault magnitude f_i. The detailed expressions for $F_W(w)$ and $f_W(w)$, as well as the evaluation of Eq. 5.22, are provided in Annex. A.2.3.

Fig. 5.4(b) illustrates the relationship between the MD probability and two key parameters: the detection interval Δ_d and fault magnitude f_i, both in the unit of σ. Larger faults tend to be detected with a lower failure rate, while wider intervals are more likely to result in MD.

Alternatively, setting the required $P(\text{MD})$ in Eq. 5.22 establishes an equation of the fault magnitude f_i. The solution can be regarded as an equivalent of the Minimum Detectable Bias (MDB) in test theory (Teunissen, 2006), ensuring that the detection failure rate against greater faults does not exceed the specified requirement. Notably, the MDB is not necessarily the worst case in the sense of integrity risk (IR), similar to the conceptions in RAIM. This aspect is further elaborated in the subsequent sections.

Comparison to RAIM detectors

To develop a common ground for comparison, this section concentrates on the detection capability, which may also be interpreted as the sensitivity to fault magnitudes.

First, prior probabilities for each hypothesis are assumed, as depicted in Tab. 5.1.

Table 5.1: Prior probabilities of multiple hypotheses for the benchmark example.

Hypothesis	Prior probability
H_1: $\mathbf{f} = [f_1,\ 0,\ 0]^T$	$P(H_1) = 10^{-3}$
H_2: $\mathbf{f} = [0,\ f_2,\ 0]^T$	$P(H_2) = 10^{-3}$
H_3: $\mathbf{f} = [0,\ 0,\ f_3]^T$	$P(H_3) = 10^{-3}$
H_0: $\mathbf{f} = [0,\ 0,\ 0]^T$	$P(H_0) = 1 - \sum_{i=1}^{3} P(H_i) = 0.997$

While being parameterized differently, the three detectors to be compared should be constructed based on equivalent conditions, e.g., the same required FA probability. In this study, the CR budget allocated to FA is set to $C_{REQ,0} = 10^{-6}$ (see Sec. 2.4.4). Subsequently, the FA probability must not be greater than

$$P(\text{FA}) = \frac{C_{REQ,0}}{P(H_0)} = 1.003 \times 10^{-6}. \tag{5.23}$$

- The **set-based detector** is characterized by the detection interval $[-\Delta_d, \Delta_d]$ intended for the set-emptiness check. Δ_d is determined by solving Eq. 5.21 as discussed earlier;
- The **RB detector** is characterized by a single test statistic q_{RB}. The test threshold T_{RB} is determined using Eq. 2.56 to Eq. 2.57;
- For the **SS detector**, the FA probability requirement is equally allocated to three test statistics q_{SS}, which is optimal for the benchmark problem. Then, Eq. 2.58 to Eq. 2.60 can be used to determine the test thresholds T_{SS}.

The obtained parameters for the three detectors are indicated in Tab. 5.2.

In the presence of faults, the distributional properties of test statistics have been fully understood:

- The RB and SS test statistics q_{RB}^2 and q_{SS}^2 follow non-central Chi-square distributions with $n - m$ and 1 degree of freedom, cf. Eq. 2.42 and Eq. 2.54, respectively;
- The analytical expressions concerning the "test statistic" W for the set-based detector are derived using *independent nonidentically distributed (inid)* order statistics, with de-

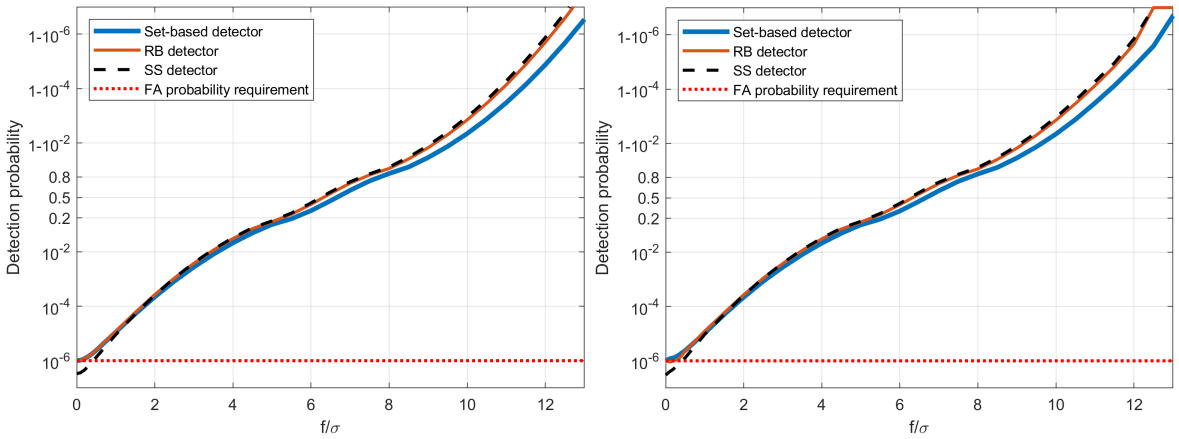
tailed formulations provided in Annex. A.2.2. Given a measurement-level fault vector $\mathbf{f}$, the detection probability is computed from respective CDF.

Table 5.2: Parameters for three detectors used in the benchmark example.

Detection method	Parameter
RB detection	$T_{RB} = 5.2560\,\sigma$
SS detection	$T_{SS} = 5.1030\,\sigma$
Set-based detection	$\Delta_d = 3.6080\,\sigma$

To verify the analytical evaluations, Monte Carlo simulations are performed by adding a constant bias f to one of the measurements in each trial, constituting the fault term $\mathbf{f}$. Consistently, the measurements are randomly generated based on standard normal distributions. Then, the detection occurrence, i.e. $P(\mathrm{D} \mid \mathbf{f})$, is counted. By varying f, the detection sensitivity is observed. Specifically, the SS detector is evaluated against a single test statistic $q_{SS,i}$ corresponding to the known fault mode. For example, $q_{SS,1}$ should be selected if the fault has been added to the first measurement.

Fig. 5.5 compares the detection probabilities for RB, SS, and set-based detectors.



(a) Analytically evaluated probabilities (b) Monte Carlo evaluated probabilities

Figure 5.5: Detection probabilities with known added faults for RB, SS and set-based detectors determined by (a) analytical evaluation, (b) Monte Carlo evaluation.

All curves start at 1.003×10^{-6} (set-based and RB detectors), or $\frac{1}{3} \times 1.003 \times 10^{-6}$ (SS detector), representing the fault-free condition. Fig. 5.5(a) and Fig. 5.5(b) show similar patterns as expected. Additionally, the comparison verifies the slightly superior performance of SS detector over RB detector (higher detection probabilities achieved). A theoretical analysis for this effect can be found in Joerger and Pervan (2014). Meanwhile, the set-based detector performs relatively worse than the other two detectors, particularly for greater biases. It may, therefore, be concluded at this stage that the classical RAIM detectors, especially the SS detector, are optimal in terms of detection capability under the Gaussian error assumption.

5.1.4 Investigation on two-dimensional scenarios

Besides observational errors, geometry is an additional important factor influencing positioning performance in higher-dimensional problems compared to the scalar-state scenario. It refers to the geometry of the lines of positioning associated with the measurements. For GNSS, it

is the constellation of the group of satellites from which signals are being received, informing (i) how many satellites are observable, (ii) how high they are in the sky, and (iii) the bearing towards them.

To further explore the potential of the set-based detector, this section examines its behavior in a two-dimensional estimation problem using Monte Carlo simulations, particularly quantifying the impact of geometry on the detector's performance. In the remainder of this subsection, an alternative geometry measure is first designed, followed by a Monte Carlo evaluation to examine the relation between the detection interval and the designed geometry measure.

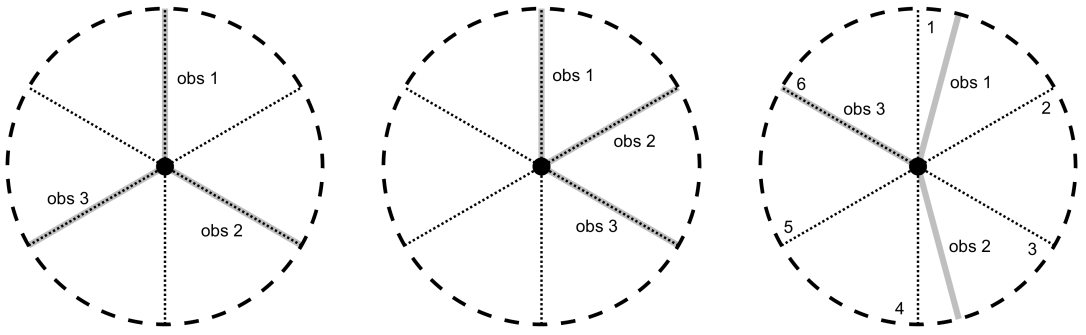
Geometry measures. The dilution of precision (DOP) is a widely-used indicator of the quality of the geometry, computed as the trace of $(\mathbf{A}^T\mathbf{A})^{-1}$ (Langley et al., 1999). While DOP is prominent for methods that utilize the least-squares estimator, its effectiveness for the set-based approach is still unclear.

To explore an alternative for DOP, a heuristic measure of two-dimensional geometry m_G is designed for the set-based detector, evaluating how far a given geometry is from an 'ideal" configuration:

$$m_G \coloneqq \min_{\{\phi_1^*,\ldots,\phi_n^*\}\subset\Psi^*} \sum_{i=1}^{n} (\phi_i - \phi_i^*)^2 \,, \tag{5.24}$$

where ϕ_i represents the azimuth angle of the i-th measurement, and ϕ_i^* belongs to a set of angles Ψ^* that can offer "ideal" configurations of geometry.

By design, under an "ideal" geometry, the set-based method should exhibit equal detection capabilities against any possible single-measurement fault. For instance, the MD probability against a given fault magnitude remains unchanged, no matter whether it affects the first, second, or third measurement. In this regard, the geometry is considered ideal when lines of positioning associated with the measurements are equally distributed within the unit circle (shown in Fig. 5.6(a)). Fig. 5.6(b) is also ideal since its positioning lines lie over the dotted lines too. Notably, the two scenarios indicate equal (and minimal) DOP values. This means that the configuration is not unique, and hence, m_G yields the minimum achievable value.



(a) "Ideal" geometry 1 (b) "Ideal" geometry 2 (c) Random geometry example

Figure 5.6: Examples of 2D geometry. The dotted lines are evenly distributed within the dashed unit circle. They indicate the ideal cases of three measurements where the lines of positioning (gray) are expected to lie (a&b). In other cases, they are not colinear (c).

In contrast, Fig. 5.6(c) illustrates an example of non-ideal geometry, which should be qualified by Eq. 5.24. Since multiple choices for $\phi_i^* \in \Phi^*$ are possible, e.g., Lines $\{1,2,3\}$, or Lines $\{1,3,6\}$, etc., it is practical to first convert ϕ_i to angles between $[0°, 180°]$, denoted by ϕ_i', and set ϕ_i^* as a sequence of $\phi_0, \phi_0 + \frac{180°}{n}, ..., \phi_0 + \frac{(n-1)\cdot 180°}{n}$. The unknown constant ϕ_0 can be

calculated by

$$\phi_0 = \arg \min_{\{\phi_1^*,\ldots,\phi_n^*\}\subset\phi^*} \sum_{i=1}^{n} (\phi_i - \phi_i^*)^2 \,, \text{ with } \phi_i^* = \phi_0 + \frac{(i-1)\cdot 180^\circ}{n}, \ \phi_i' \in [0, 180]^\circ. \quad (5.25)$$

Subsequently, ϕ_i^* is obtained, and m_G can be determined using Eq. 5.24.

Detection interval evaluation. The detection interval $[-\Delta_d, \Delta_d]$ associated with a specific FA probability is used to characterize the set-based detector. In the Monte Carlo context, $P(\text{FA})$ corresponds to the occurrence of an empty polytope $\mathcal{P}_d$ under nominal (fault-free) conditions (cf. Eq. 5.8).

For a given geometry, normally distributed measurement samples are generated for 1000 epochs, with 3 to 15 measurements per epoch, simulating nominal conditions. For each trial (a fixed number of measurements at a single epoch), the set-based detector is applied with a given Δ_d. The null hypothesis is either correctly accepted ($\mathcal{P}_d \neq \emptyset$) or wrongly rejected ($\mathcal{P}_d = \emptyset$, FA event). The objective is to determine a proper Δ_d to ensure that the FA occurrence does not exceed the required probability.

Starting from an initial guess, Δ_d is iteratively refined until the desired FA probability is achieved, which is set to $P(\text{FA}) = 10^{-2}$ in this study. Once Δ_d is determined, the set-based detector is considered calibrated to guarantee the required FA probability for users under the given geometry.

Geometry measures evaluation. A specific geometry is represented by randomly generated azimuth angles (uniformly distributed within $[0, 2\pi]$). For a given number of measurements, 2000 various geometries are generated. They are measured by either DOP or the designed parameter m_G.

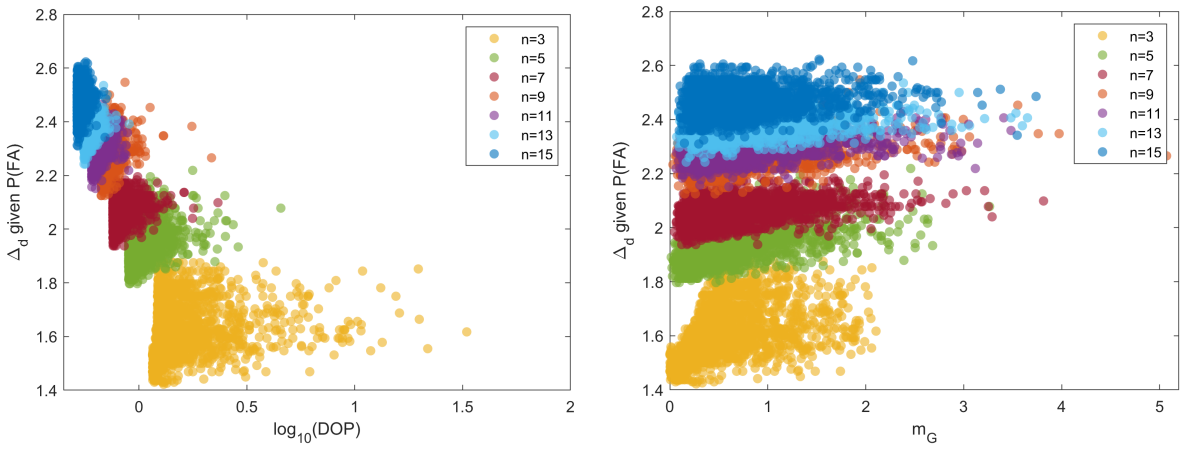
Fig. 5.7 depicts the relation between the geometry measures (DOP or m_G) and the detection interval (Δ_d/σ). DOP can capture the impact of measurement numbers, i.e., more measurements tend to reduce the DOP value, which results in a lower chance of FA for a given Δ_d. The designed measure m_G, on the contrary, does not present a functional relationship with the measurement numbers. Nevertheless, it can capture the quality of geometry, which is more significant with fewer measurements. This trend is visible in the density plots in Fig. 5.8.

In conclusion, the DOP is not an effective indicator of geometrical quality for the set-based detector as it is in stochastic approaches. A heuristic measure m_G has been developed to serve two main purposes: (i) to indicate the detection capability of the set-based detector under specific conditions and (ii) to assist in determining detection intervals that meet certain requirements. However, further improvements to this measure are necessary to address higher-dimensional problems.

5.1.5 Enhancement of robustness with weighting models

In urban navigation scenarios, potential extra path delay (EPD) due to signal reflections and refraction can severely bias the GNSS measurements and, subsequently, threaten the navigation quality. These threats are considered unbounded faults, and they are expected to be detected and excluded before a navigation solution is estimated without degrading system integrity and availability.

Sec. 5.1.3 reveals that the fault magnitude strongly influences the detection capability of the set-based method. To enhance the detector's robustness, weighting models may be applied.



(a) DOP as geometry measure

(b) Designed heuristic geometry measure

Figure 5.7: Impact of geometry on the set-based detector in 2D scenarios. Two geometry measures (DOP and a heuristic measure m_G) are explored, and their relation with detection intervals Δ_d is demonstrated.

If any external information indicates that a subset of measurements is very likely biased but remains undetected under the current settings, reducing their detection intervals can help increase sensitivity to undetectable biases.

The vector of detection intervals can be adjusted by multiplying it with a weighting model $\mathbf{T}$ in Eq. 5.6:

$$\begin{cases} \mathbf{A}\mathbf{x} \leq \mathbf{y} + \mathbf{T}\boldsymbol{\Delta}_d \\ -\mathbf{A}\mathbf{x} \leq -\mathbf{y} + \mathbf{T}\boldsymbol{\Delta}_d \end{cases}. \tag{5.26}$$

Inspired by the weighted least-squares estimator, this thesis proposes a method for determining the matrix $\mathbf{T}$ by decomposing the weighting matrix $\mathbf{P}$ that is used in stochastic approaches. The desired weighting matrix is given by

$$\mathbf{P} = (\mathbf{T})^T\mathbf{T}, \tag{5.27}$$

which is a diagonal matrix. For example, $\mathbf{T}$ could represent either an elevation-dependent model (EDM) or a signal strength-dependent model (SDM). In the case of EDM, the diagonal matrix is constructed as follows:

$$\mathbf{T}_{\text{EDM}} = \text{diag}\left(\sqrt{\sin\theta}\right), \tag{5.28}$$

where θ is a vector of elevation angles expressed in radians.

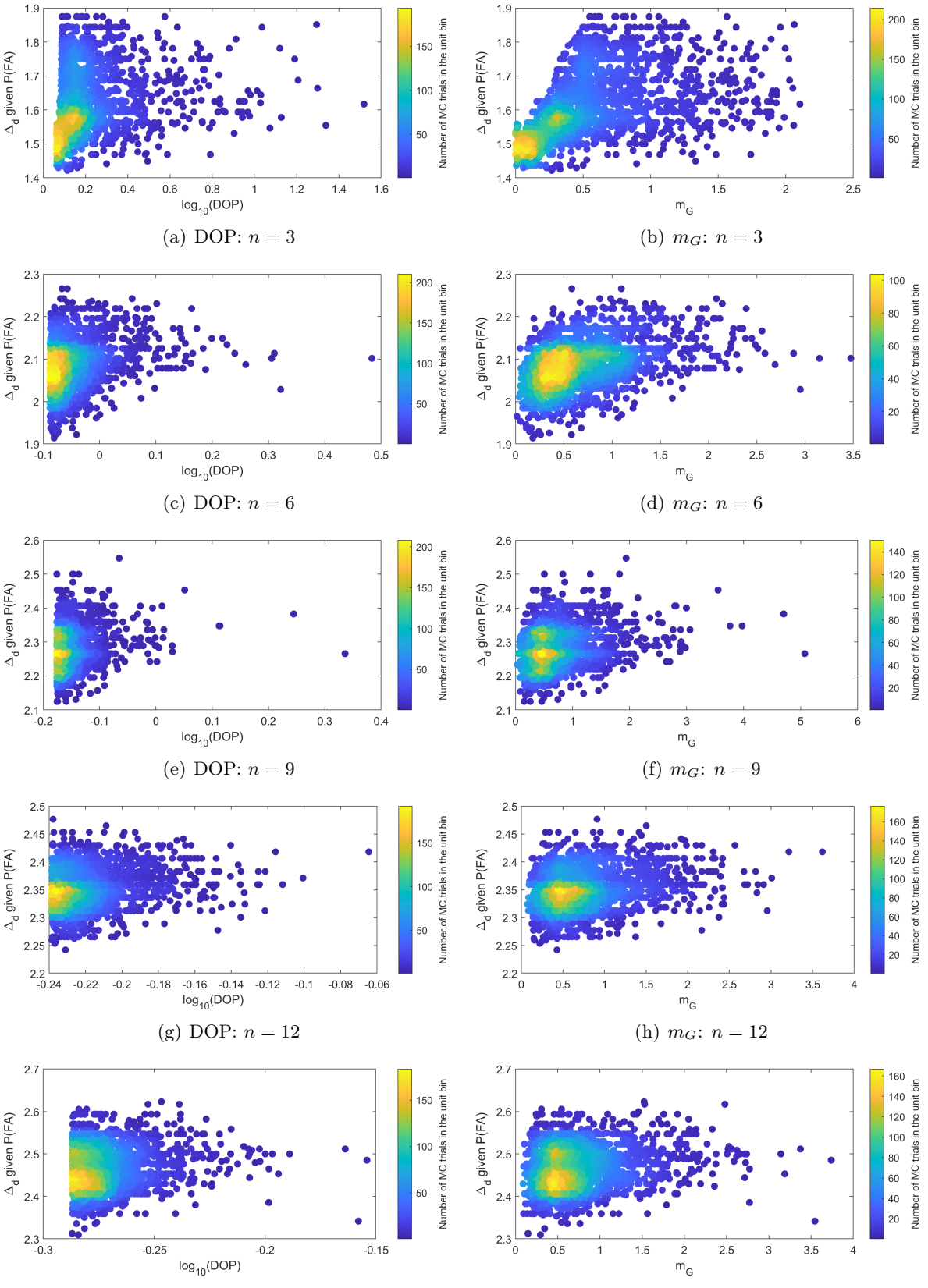
For SDM, a possible diagonal matrix may be constructed based on the *sigma-ε* model (Hartinger and Brunner, 1999), which determines the variance of GNSS measurements using C/N0 information:

$$\sigma_i^2 = V_i + C_i \cdot 10^{-\frac{\text{C/N0}}{10}}, \tag{5.29}$$

where V_i and C_i are model parameters in the unit of m^2 and $\text{m}^2 \cdot \text{Hz}$. C/N0 is the measured carrier-to-noise power-density ratio in the unit of dB-Hz. Subsequently, the proposed weighting matrix reads

$$\mathbf{T}_{\text{SDM}} = \text{diag}\left(\mathbf{t}_{\text{SDM}}\right), \text{ with } t_{\text{SDM},i} = \sqrt{V_i + C_i \cdot 10^{-\frac{\text{C/N0}}{10}}}. \tag{5.30}$$

The above examples are not exclusive and may be adapted according to the measurement environment, hardware, etc.



(a) DOP: $n = 3$ (b) m_G: $n = 3$

(c) DOP: $n = 6$ (d) m_G: $n = 6$

(e) DOP: $n = 9$ (f) m_G: $n = 9$

(g) DOP: $n = 12$ (h) m_G: $n = 12$

(i) DOP: $n = 15$ (j) m_G: $n = 15$

Figure 5.8: Impact of geometry on the set-based detector with various numbers of measurements in 2D scenarios. Two geometry measures (DOP and a heuristic measure m_G) are explored, and their relation with detection intervals Δ_d is highlighted.

5.2 Development of interval-based integrity monitoring

This section focuses on interval-based autonomous integrity monitoring for the model presented in Eq. 5.1. The proposed approach employs set-based fault detection and exclusion (FDE) functions in conjunction with the computation of Protection Level (PL) that is delivered to users. The development begins with the definition of the loss of integrity (LOI) in Sec. 5.2.1. This is followed by an analytical evaluation of integrity risk (IR) bound and the design of an optimal estimator aimed at minimizing the PL. Additionally, the strategy's architecture is outlined, and a sub-optimal solution is introduced to address practical requirements.

5.2.1 Evaluation for loss of integrity

Definition of loss of integrity

The classical definition of the Hazardous Misleading Information (HMI) or loss of integrity (LOI) in RAIM refers to the concurrent occurrence of *No Detection (ND)* and *Position Failure (PF)* (see Eq. 2.61):

- ***No Detection***: The detection is not triggered, implying that the measurements are presumed self-consistent. In the context of set-based detection, ND occurs when the detection polytope ($\mathcal{P}_d$) remains non-empty.
- ***Position Failure***: The Position Error (PE), denoted by ε, exceeds the Alert Limit (AL, ℓ), indicating that the Euclidean distance between the estimated position ($\hat{\mathbf{x}}$) and the true position ($\mathbf{x}^*$) surpasses ℓ, i.e., $\|\varepsilon\| := \|\hat{\mathbf{x}} - \mathbf{x}^*\| > \ell$.

The HMI probability, also referred to as integrity risk (IR), is then formulated as follows:

$$\begin{aligned} P(\text{HMI}) &= P\left(\text{ND} \wedge \text{PF}\right) \\ &= P\left(\mathcal{P}_d \neq \emptyset \wedge \|\varepsilon\| > \ell\right) \leq I_{REQ}, \end{aligned} \tag{5.31}$$

where I_{REQ} represents the IR requirement.

Integrity risk bound derivation

Starting from the definition in Eq. 5.31, it is feasible to establish an upper bound as follows:

$$P(\text{HMI}) = P(\text{ND} \wedge \text{PF}) = P(\text{ND} \mid \text{PF}) \cdot P(\text{PF}) \leq P(\text{PF}). \tag{5.32}$$

Here, the PF probability is used to upper bound the HMI probability since $P(\text{ND} \mid \text{PF}) \leq 1$ hold true regardless of the dependence between ND and PF. This approach may introduce some looseness to the resulting IR bound, which is further discussed in the subsequent analyses. Nevertheless, it offers the following benefits:

- It protects the system in any case of missed detection;
- It alleviates the need for analytically quantifying test thresholds in fault detection;
- It supports the entire FDE functions, ensuring protection even if a fault cannot be identified or excluded;
- It allows the integration of any FD method into the system without harming its integrity.

Similar concepts are adopted in the literature, e.g., Bayesian RAIM, a variant of RAIM that takes measurements as inputs, defines Eq. 5.32 as posterior integrity risk.

The next step is to evaluate $P(\mathrm{PF})$. The interval-based integrity monitor takes advantage of a *feasible solution set* (FSS, denoted by $\mathcal{S}$) in the state domain to determine the point estimate and assess the integrity risk. As described in Chap. 4, $\mathcal{S}$ indicates the possible range for the truth with high confidence. This setup introduces the probability of a *No Protection (NP)* event:

$$P(\mathrm{NP}) = P(\mathbf{x}^* \notin \mathcal{S}), \tag{5.33}$$

which quantifies the probability that $\mathcal{S}$ fails to enclose the truth.

To establish the upper bound, a functional model $g(\mathcal{S})$, yet unspecified, is needed to serve as a scalar measure of $\mathcal{S}$. This model must ensure that the estimation error ε satisfies the following condition:

$$\|\varepsilon\| := \|\hat{\mathbf{x}} - \mathbf{x}^*\| \leq g(\mathcal{S}), \tag{5.34}$$

for all cases where $\mathbf{x}^* \in \mathcal{S}$. In Chap. 4, the term $\|\varepsilon\|$ is defined as the estimation error bound for a set-based point estimator.

$g(\mathcal{S})$ should admit two critical properties:

- It reflects the geometric configuration of $\mathcal{S}$ and is **unique** with respect to $\mathcal{S}$;
- It is **scalable**, meaning that adjustments to $\mathcal{S}$ proportionally alter $g(\mathcal{S})$.

By scaling $\mathcal{S}$ to guarantee $g(\mathcal{S}) \leq \ell$, the corresponding probabilistic assurance also varies. Therefore, an inequality for $P(\mathrm{PF})$ can be sought:

$$\begin{aligned} P(\mathrm{PF}) :=& P\left(\|\varepsilon\| > \ell\right) \\ \leq& P\left(\|\varepsilon\| > g(\mathcal{S}) \mid g(\mathcal{S}) \leq \ell\right). \end{aligned} \tag{5.35}$$

Taking Eq. 5.33 into consideration, Eq. 5.35 can be broken down into two components:

$$\begin{aligned} &P\left(\|\varepsilon\| > g(\mathcal{S}) \mid g(\mathcal{S}) \leq \ell\right) = \\ &\quad P\Big(\left(\|\varepsilon\| > g(\mathcal{S}) \wedge \mathbf{x}^* \in \mathcal{S}\right) \Big| g(\mathcal{S}) \leq \ell\Big) + P\Big(\left(\|\varepsilon\| > g(\mathcal{S}) \wedge \mathbf{x}^* \notin \mathcal{S}\right) \Big| g(\mathcal{S}) \leq \ell\Big). \end{aligned} \tag{5.36}$$

By definition, the first component is zero because the estimation error is bounded by $g(\mathcal{S})$ when $\mathbf{x}^* \in \mathcal{S}$. The second component, representing the error when $\mathbf{x}^* \notin \mathcal{S}$, is further bounded through the inequalities in Eq. 5.37. For better readability, $\|\varepsilon\| > g(\mathcal{S})$ is denoted by *Event A*, $\mathbf{x}^* \notin \mathcal{S}$ by *Event B*, and $g(\mathcal{S}) \leq \ell$ by *Event C*. Hence,

$$\begin{aligned} P(AB \mid C) =& \frac{P(ABC)}{P(C)} = \frac{P(BC) - P(\overline{A}BC)}{P(C)} = P(B \mid C) - P(\overline{A}B \mid C) \\ \leq& P(B \mid C), \end{aligned} \tag{5.37}$$

where $\overline{A}$ is the complement of *Event A*.

These inequalities lead to an upper bound for $P(\mathrm{PF})$:

$$P\left(\|\varepsilon\| > \ell\right) \leq P\left(\mathbf{x}^* \notin \mathcal{S} \mid g(\mathcal{S}) \leq \ell\right), \tag{5.38}$$

which is a conditional NP probability. It is obtained by disregarding the concurrence of $\|\varepsilon\| \leq g(\mathcal{S})$ and $\mathbf{x}^* \notin \mathcal{S}$. The potential conservatism is systematic because a scalar measure $g(\mathcal{S})$ is never sufficient to fully capture the geometrical information of $\mathcal{S}$, especially in scenarios

involving higher dimensions or complex geometries. This aspect is further analyzed with illustrative examples in the remainder.

Conclusively, the final IR bound is expressed as

$$\begin{aligned} P(\mathrm{ND} \wedge \mathrm{PF}) &\leq P(\mathrm{NP} \mid g(\mathcal{S}) \leq \mathrm{AL}), \\ \text{i.e., } P\left(\mathcal{P}_d \neq \emptyset \wedge \|\varepsilon\| > \ell\right) &\leq P(\mathbf{x}^* \notin \mathcal{S} \mid g(\mathcal{S}) \leq \ell). \end{aligned} \tag{5.39}$$

5.2.2 Integrity monitoring against multiple faults

In the presence of multiple simultaneous faults, which are likely to occur with multi-GNSS in urban scenarios (as reported by Ruwisch and Schön (2024)), it is crucial to enhance the integrity performance against multiple faults. To achieve this, the interval-based approach adopts a multiple hypothesis framework, which is used in ARAIM. Under this framework, multiple hypotheses are considered for each potential fault mode. Each hypothesis H_i for $i = 0, 1, ...h$ assumes that a certain subset of measurements is faulty while the rest are nominal (fault-free). Specifically, the null hypothesis H_0 represents the case where all measurements are fault-free.

A key advantage of the multiple hypothesis framework is that it establishes a straightforward link between the threat model, PL and integrity risk (Blanch et al., 2007). It provides explicit probabilistic interpretations for the interval-based integrity monitor, which in turn enables effective bounding of unscheduled and undetected faults down to very low probabilities.

Equation for Protection Level

In practical applications, it is more computationally efficient to calculate the position-domain Protection Level (PL) based on required IR rather than evaluating IR from a specified AL. By referencing the computed PL to AL, the user will be informed whether the navigation integrity is assured under the current satellite geometry and availability.

The baseline ARAIM algorithm achieves this by adapting the IR bound formula (replacing AL with PL) and then solving the obtained equation. Similarly, concerning the established IR bound for the interval approach in Eq. 5.39, the desired PL equation is formulated as

$$P(\mathbf{x}^* \notin \mathcal{S} \mid g(\mathcal{S}) = \mathrm{PL}) = I_{REQ}. \tag{5.40}$$

Eq. 5.40 is not directly solvable, which necessitates a two-step solution:

1. **Determine the feasible solution set $\mathcal{S}$.** Limit the NP probability with respect to $\mathcal{S}$ to meet the IR requirement using

$$P(\mathbf{x}^* \notin \mathcal{S}) = I_{REQ}. \tag{5.41}$$

2. **Compute the scalar measure $g(\mathcal{S})$ as PL**. Determine the PL based on the designed scalar measure:

$$\mathrm{PL} = g(\mathcal{S}). \tag{5.42}$$

In Step 1, the fundamental elements for constructing $\mathcal{S}$ have been fully defined, though the procedures remain unspecified at this stage. Known satellite geometry from the linearized model (Eq. 4.2) and measurement intervals derived from observational uncertainty models support this process.

In Step 2, the PL is computed through the functional model $g(\mathcal{S})$, ensuring adherence to the derived IR bound.

Now it becomes clear that the keys to this strategy include the determination of the FSS $\mathcal{S}$, a proper scalar measure $g(\mathcal{S})$ and the corresponding point estimator $\hat{\mathbf{x}}_{\mathcal{S}}$. These aspects are thoroughly addressed in the remainder of this section.

In general, the above strategy is straightforward and offers two significant advantages:

- The PL and IR are linked through the construction of FSS.
- By design, the computed PL is not smaller than PE at required probabilities, ensuring that the PL provides a reliable and conservative estimate of the position uncertainty.

Feasible solution set for integrity purpose

Based on the principles discussed in Sec. 4.3.1, a feasible solution set (FSS) can be established using the set estimator based on constraint satisfaction. For the fault-free case (H_0) in the model of Eq. 5.1, a polytope solution set $\mathcal{P}_{s,0}$ can be obtained as follows:

$$\mathcal{P}_{s,0} = \left\{ \mathbf{x} \in \mathbb{R}^m \mid \mathbf{B}\,\mathbf{x} \leq \mathbf{b}, \mathbf{B} = \begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \end{bmatrix} \in \mathbb{R}^{2n \times m}, \mathbf{b} = \begin{bmatrix} \mathbf{y} - \underline{\boldsymbol{\Delta}}_s \\ -\mathbf{y} + \overline{\boldsymbol{\Delta}}_s \end{bmatrix} \in \mathbb{R}^{2n} \right\}. \tag{5.43}$$

Here, the measurement intervals are denoted by $[\boldsymbol{\Delta}_s] = [\underline{\boldsymbol{\Delta}}_s, \overline{\boldsymbol{\Delta}}s]$.

In the multiple hypothesis framework, the FSS $\mathcal{S}$ must account for potential unbounded or undetected faults, which $\mathcal{P}_{s,0}$ does not fully capture. To address this, $\mathcal{S}$ is defined as the union of multiple polytope solution sets:

$$\mathcal{S} = \bigcup_{i=0}^{h} \mathcal{P}_{s,i}, \tag{5.44}$$

where $\mathcal{P}_{s,i}$ corresponds to the polytope solution set for each hypothesis H_i. For instance, consider H_i for $i > 0$, where the measurement y_i is assumed to be faulty. However, no information is available regarding the fault magnitude f_i, nor should any assumption be made. By expanding the interval bounds for the corresponding measurements to infinity, the resulting solution set can tolerate the impact of potential faults on estimation. The inequalities can be expressed as

$$\begin{cases} \mathbf{A}_i\mathbf{x} \leq {\mathbf{H}_i}^T\mathbf{y} + {\mathbf{H}_i}^T\boldsymbol{\Delta}_\mathbf{s}, \\ \mathbf{A}_i\mathbf{x} \leq \mathbf{H}_i^{*T}\mathbf{y} + \mathbf{H}_i^{*T}\infty, \\ -\mathbf{A}_i\mathbf{x} \leq -{\mathbf{H}_i}^T\mathbf{y} + {\mathbf{H}_i}^T\boldsymbol{\Delta}_\mathbf{s}, \\ -\mathbf{A}_i\mathbf{x} \leq -\mathbf{H}_i^{*T}\mathbf{y} + \mathbf{H}_i^{*T}\infty, \end{cases} \tag{5.45}$$

where $\mathbf{H}_i$, $\mathbf{H}_i^*$, and $\mathbf{A}_i$ are defined in Eq. 2.48 and Eq. 2.49.

$\mathcal{P}_{s,i}$ for $i > 0$ is designed to mitigate the influence of specific faulty measurements, making it a fault-tolerant solution set with respect to the i-th fault mode. Therefore, the union $\mathcal{S}$ can be interpreted as a solution set that tolerates the impact of any potential faults across all monitored fault modes. Additionally, Eq. 5.45 can be reformulated into a shorter version:

$$\begin{cases} \mathbf{A}_i\mathbf{x} \leq {\mathbf{H}_i}^T\mathbf{y} + {\mathbf{H}_i}^T\boldsymbol{\Delta}_\mathbf{s}, \\ -\mathbf{A}_i\mathbf{x} \leq -{\mathbf{H}_i}^T\mathbf{y} + {\mathbf{H}_i}^T\boldsymbol{\Delta}_\mathbf{s}. \end{cases} \tag{5.46}$$

This simplification excludes assumed faulty measurements and utilizes the remaining fault-free ones. It is mathematically equivalent to Eq. 5.45 and allows $\mathcal{P}_{s,i}$ to be expressed as

$$\mathcal{P}_{s,i} = \left\{ \mathbf{x} \in \mathbb{R}^m \mid \mathbf{B}_i\,\mathbf{x} \leq \mathbf{b}_i, \mathbf{B}_i = \begin{bmatrix} \mathbf{A}_i \\ -\mathbf{A}_i \end{bmatrix} \in \mathbb{R}^{2n \times m}, \mathbf{b}_i = \begin{bmatrix} {\mathbf{H}_i}^T(\mathbf{y} - \underline{\boldsymbol{\Delta}}_s) \\ -{\mathbf{H}_i}^T(\mathbf{y} - \overline{\boldsymbol{\Delta}}_s) \end{bmatrix} \in \mathbb{R}^{2n} \right\}. \tag{5.47}$$

It is useful to note the hierarchical relationship among different $\mathcal{P}_{s,i}$. For example, consider any two hypotheses H_{i_1} and H_{i_2}. If the faulty measurements under H_{i_1} are a subset of those under H_{i_2}, then $\mathcal{P}_{s,i_1} \subseteq \mathcal{P}_{s,i_2}$. This inclusion relationship reveals the construction of $\mathcal{S}$: The active constraints defining the boundaries of $\mathcal{S}$ are predominantly shaped by the polytopes with the fewest fault-free measurements (see Fig. 5.9). It also underscores the adaptability of FSS to varying levels of system integrity, ensuring that the most critical scenarios dictate the safety margins.

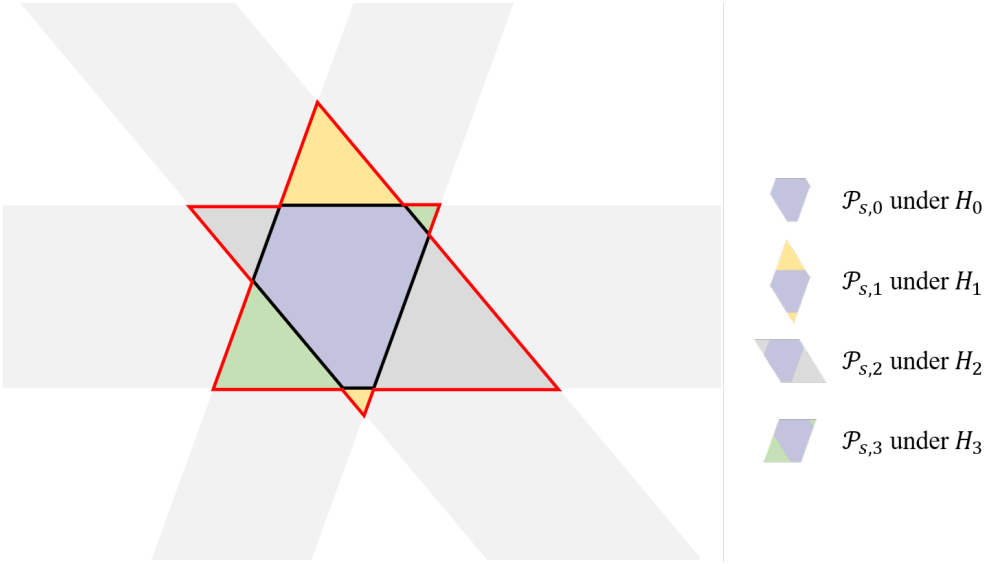


Figure 5.9: Conceptual sketch for the feasible solution set $\mathcal{S}$ in a two-dimensional example with three measurements (slabs): (i) with $r = 0$, $\mathcal{P}_{s,0}$ configures $\mathcal{S}$ in black contour; (ii) with $r = 1$, the envelope of $\mathcal{P}_{s,1}$, $\mathcal{P}_{s,2}$ and $\mathcal{P}_{s,3}$ defines $\mathcal{S}$ in red contour.

Evaluation for No Protection probability

In this analysis, a *bounding probability* for any fault-free measurement interval is referred to as P_{Δ_s}, representing the probability that the measurement error is bounded by an interval $[-\Delta_s, \Delta_s]$. The alternative interpretation, as in Eq. 4.24, is also valid, i.e., the probability that the expected observation y^* is contained by the interval $[y - \Delta_s, y + \Delta_s]$. Utilizing the binomial law, it is possible to calculate relevant probabilities for each fault-tolerant polytope $\mathcal{P}_{s,i}$.

Under H_0: The probability of $\mathcal{P}_{s,0}$ containing the true state is simply expressed as

$$P(\mathbf{x}^* \in \mathcal{P}_{s,0} \mid H_0) = (P_{\Delta_s})^n. \tag{5.48}$$

Reflecting on $\mathcal{S}$ as the union of $h+1$ polytopes, the probability that $\mathcal{S}$ provides protection under H_0 is computed by summing up the probabilities across all polytopes. For instance, the polytope $\mathcal{P}_{S,j}$ is formed by $n - j$ measurement intervals:

$$P(\mathbf{x}^* \in \mathcal{P}_{S,j} \mid H_0) = (P_{\Delta_s})^{n-j} \cdot (1 - P_{\Delta_s})^j, \tag{5.49}$$

which can be interpreted as a polytope solution set that tolerates any potential errors in the j measurements. And,

$$\mathcal{P}_{s,0} \subseteq \mathcal{P}_{s,j}, \text{ for } j > 0. \tag{5.50}$$

Hence, the conditional NP probability for $\mathcal{S}$ under H_0 is

$$P(\mathbf{x}^* \notin \mathcal{S} \mid H_0) = 1 - \sum_{j=0}^{r} \binom{r}{j} \cdot (P_{\Delta_s})^{n-j} \cdot (1 - P_{\Delta_s})^j, \tag{5.51}$$

where r is the maximum number of tolerated measurements, which must be specified in advance.

It is worth noting that Eq. 5.51 has been used in various prior works, although often with different interpretations. For example, Drevelle and Bonnifait (2009) applies a similar formulation without considering a multiple-hypothesis framework, leading to differing implications for integrity monitoring.

Under H_i: For an alternative hypothesis H_i where n_i out of n measurements are deemed faulty and $n - n_i$ are fault-free, the probability that the fault-tolerant polytope $\mathcal{P}_{s,i}$ encloses the truth is computed by

$$P(\mathbf{x}^* \in \mathcal{P}_{s,i} \mid H_i) = (P_{\Delta_s})^{n-n_i}. \tag{5.52}$$

Excluding n_i faulty measurements allows for the omission of assumptions on the distribution or magnitude of the faults. This approach ensures that the induced position error remains bounded, given available information.

Similarly, if a maximum of r out of n measurements are monitored by $\mathcal{S}$, a tolerance of up to $r - n_i$ measurements is permitted for any potential errors under H_i ($n_i \leq r$). The conditional NP probability for $\mathcal{S}$ under H_i is expressed as

$$P(\mathbf{x}^* \notin \mathcal{S} \mid H_i) = 1 - \sum_{j=0}^{r-n_i} \binom{r-n_i}{j} \cdot (P_{\Delta_s})^{n-n_i-j} \cdot (1 - P_{\Delta_s})^j. \tag{5.53}$$

Therefore, the overall NP probability in the multiple-hypothesis framework integrates the conditional NP probabilities for all hypotheses weighted by their prior probabilities $P(H_i)$:

$$\begin{aligned} P\left(\mathbf{x}^* \notin \mathcal{S}\right) &= \sum_{i=0}^{h} P\left(\mathbf{x}^* \notin \mathcal{S} \mid H_i\right) \cdot P(H_i) \\ &= \sum_{i=0}^{h} \left(1 - \sum_{j=0}^{r-n_i} \binom{r-n_i}{j} \cdot (P_{\Delta_s})^{n-n_i-j} \cdot (1 - P_{\Delta_s})^j\right) \cdot P(H_i). \end{aligned} \tag{5.54}$$

The ARAIM threat model can be leveraged. Relevant probabilities include:

- Single satellite failure probability: P_{sat};
- Probability of n_i simultaneous satellite failure: $P(n_i) = \binom{n}{n_i} \cdot (P_{sat})^{n_i} \cdot (1 - P_{sat})^{n-n_i}$;
- Probability of fault-free mode (H_0): $P(H_0) = P(n_i = 0) = (1 - P_{sat})^n$;
- Probability that more than r simultaneous satellite failures occur but are not monitored. It is used to upper bound their IR contribution:

$$P_{NM} = 1 - \sum_{n_i=0}^{r} P(n_i). \tag{5.55}$$

  This term is for practical usage since some fault modes with multiple simultaneous failures can rarely occur and can be ignored compared to the IR requirement. A threshold P_{sat_thres} can be pre-defined for determining r by limiting P_{NM}, i.e.,

$$1 - \sum_{n_i=0}^{r} P(n_i) \leq P_{sat\_thres}. \tag{5.56}$$

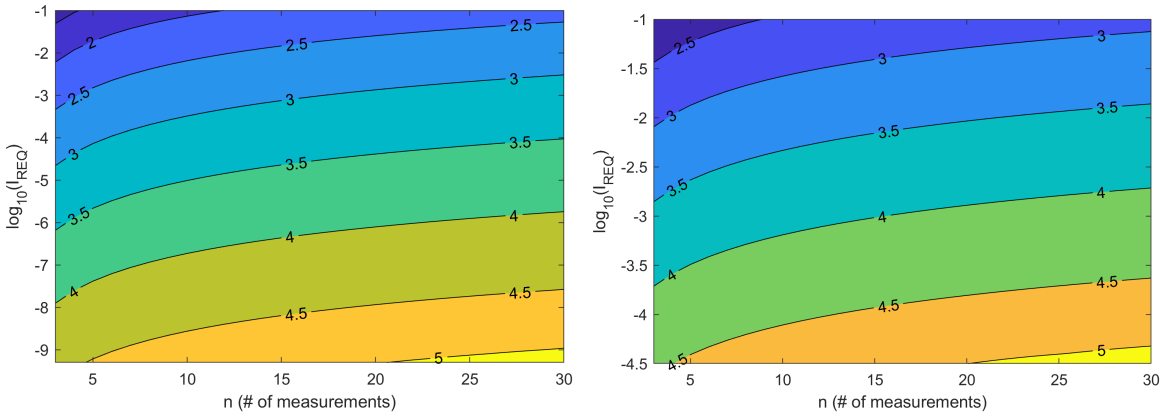
The above model is simplified without considering constellation failure probabilities. A completed description can be found in the baseline ARAIM algorithm description document (WG-C ARAIM TSG, 2019).

Accordingly, Eq. 5.54 can be further rewritten as

$$P\left(\mathbf{x}^* \notin \mathcal{S}\right) \leq \sum_{n_i=0}^{r} \left(1 - \sum_{j=0}^{r-n_i} \binom{r-n_i}{j} \cdot (P_{\Delta_s})^{n-n_j-j} \cdot (1-P_{\Delta_s})^{j}\right) \cdot P(n_i) + P_{NM}. \quad (5.57)$$

Eq. 5.57 reveals a comprehensive framework for evaluating the NP probability. By applying Eq. 5.57 to Eq. 5.41, one can solve for P_{Δ_s}, which enables the determination of $[\boldsymbol{\Delta}_s]$ based on specific observation uncertainty model. For example, if a Gaussian model is assumed, the relevant formulations in Eq. 4.26-4.28 should be used. Additionally, Tab. 4.4 offers further guidance on incorporating both stochastic and systematic uncertainty representations. Once $[\boldsymbol{\Delta}_s]$ is defined, the feasible solution set $\mathcal{S}$ is fully determined.

As an example, assuming $P_{sat} = 10^{-5}$, the interval radius (normalized by measurement error standard deviation Δ/σ) is jointly influenced by the number of measurements and IR requirements, as shown in the contour plots in Fig. 5.10(a) for the case of $r = 1$, and in Fg. 5.10(b) for $r = 2$.



(a) Δ/σ for $r = 1$ (b) Δ/σ for $r = 2$

Figure 5.10: Normalized interval radius (Δ/σ) with respect to the number of measurements and IR requirement as contour plots. Identical satellite-wise P_{sat} is assumed, and the subfigures correspond to various numbers of maximum simultaneous satellite failures monitored (r)

- More satellites (hence more chance of fault) require larger intervals to ensure the IR requirement;
- Stricter IR requirements can only be satisfied with wider intervals if the number of satellites is unchanged;
- Excluding more fault modes from being monitored must be compensated for by enlarging intervals to ensure the IR requirement.

These features are remarkable and essential to understanding the characteristics of the desired interval approach, which are further explored with demonstrative examples in the following sections.

Optimal estimator design to minimize PL

Once the FSS is determined, the remaining questions include the design of an optimal point estimator and the desired scalar measure $g(\mathcal{S})$, computed as PL.

While the PL is defined to be the position bound, it is expected that

$$\begin{aligned} \mathrm{PL} &\geq \max \|\varepsilon\|, \\ \text{i.e., } \mathrm{PL} &\geq \max_{\hat{\mathbf{x}},\mathbf{x}\in\mathbb{R}^m} \|\hat{\mathbf{x}} - \mathbf{x}^*\|. \end{aligned} \tag{5.58}$$

Eq. 5.58 may be discussed under two conditions concerning $\mathcal{S}$: $\mathbf{x}^* \in \mathcal{S}$ and $\mathbf{x}^* \notin \mathcal{S}$.

- **Condition 1 ($\mathbf{x}^* \in \mathcal{S}$).** It is desired that

$$g(\mathcal{S}) \geq \max_{\hat{\mathbf{x}}\in\mathbb{R}^m, \mathbf{x}^*\in\mathcal{S}} \|\mathbf{x}^* - \hat{\mathbf{x}}\|. \tag{5.59}$$

  Subsequently, the "*best*" achievable scalar measure can be formulated as

$$g(\mathcal{S}) = \arg\min_{g(\mathcal{S})} \max_{\mathbf{x}\in\mathbb{R}^m, \mathbf{x}^*\in\mathcal{S}} \|\mathbf{x}^* - \mathbf{x}\|, \tag{5.60}$$

  which is indeed a minimax estimation problem. Equivalently, an optimal point estimator can be formulated as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{R}^m} \max_{\mathbf{x}^*\in\mathcal{S}} \|\mathbf{x}^* - \mathbf{x}\|. \tag{5.61}$$

- **Condition 2 ($\mathbf{x}^* \notin \mathcal{S}$).** Fortunately, this case is captured by I_{REQ}, cf. Eq. 5.41, and therefore, it needs no longer to be considered.

Sec. 4.3.2 has provided a solution to problems like Eq. 5.60 and Eq. 5.61. Its optimal estimator is the Chebyshev center of $\mathcal{S}$ that minimizes the worst-case error bound ($\max \|\varepsilon\|$), which is its Chebyshev radius. Hence,

$$\hat{\mathbf{x}} = \mathrm{cen}(\mathcal{S}), \text{ and } \mathrm{PL} = \mathrm{rad}(\mathcal{S}). \tag{5.62}$$

Notably, $\mathcal{S}$ may not be a convex polytope. Nevertheless, the *existence* and *uniqueness* of $\hat{\mathbf{x}} = \mathrm{cen}(\mathcal{S}) \in \mathbb{R}^m$ still hold since $\mathcal{S}$ is formed as the union of a set of polytopes.

A conceptual example is given in Fig. 5.11. The red dashed contour represents the minimum bounding circle enclosing the union of four fault-tolerant polytopes. The center and radius of the circle indicate the point estimator $\hat{\mathbf{x}}$ and PL by definition.

The problem of the minimum bounding spheres, or the smallest enclosing circle problem, is of interest to various disciplines, such as computer graphics and computational geometry. A number of algorithms have been developed to address this problem, such as Seidel (1991), Ritter (1990), Welzl (1991), and Skyum (1991). As an example, the method in Seidel (1991) solves the linear programming algorithm in expected $\mathcal{O}(n)$ computation time.

In particular, this problem degrades to finding the minimum bounding interval for the scalar estimation problem ($m = 1$), which can be efficiently solved by a sorting algorithm. The corresponding optimal estimator is the midpoint of the minimum bounding interval. Sec. 5.3.1 further explores this feature, based on which a practical solution to integrity monitoring for road applications is developed in Sec. 5.2.4.

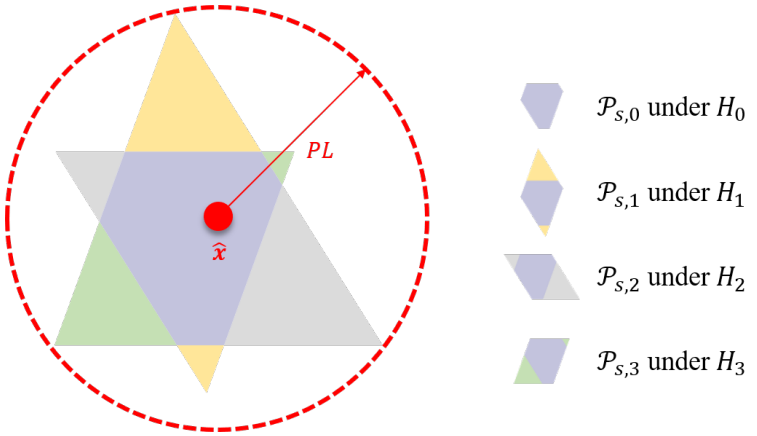
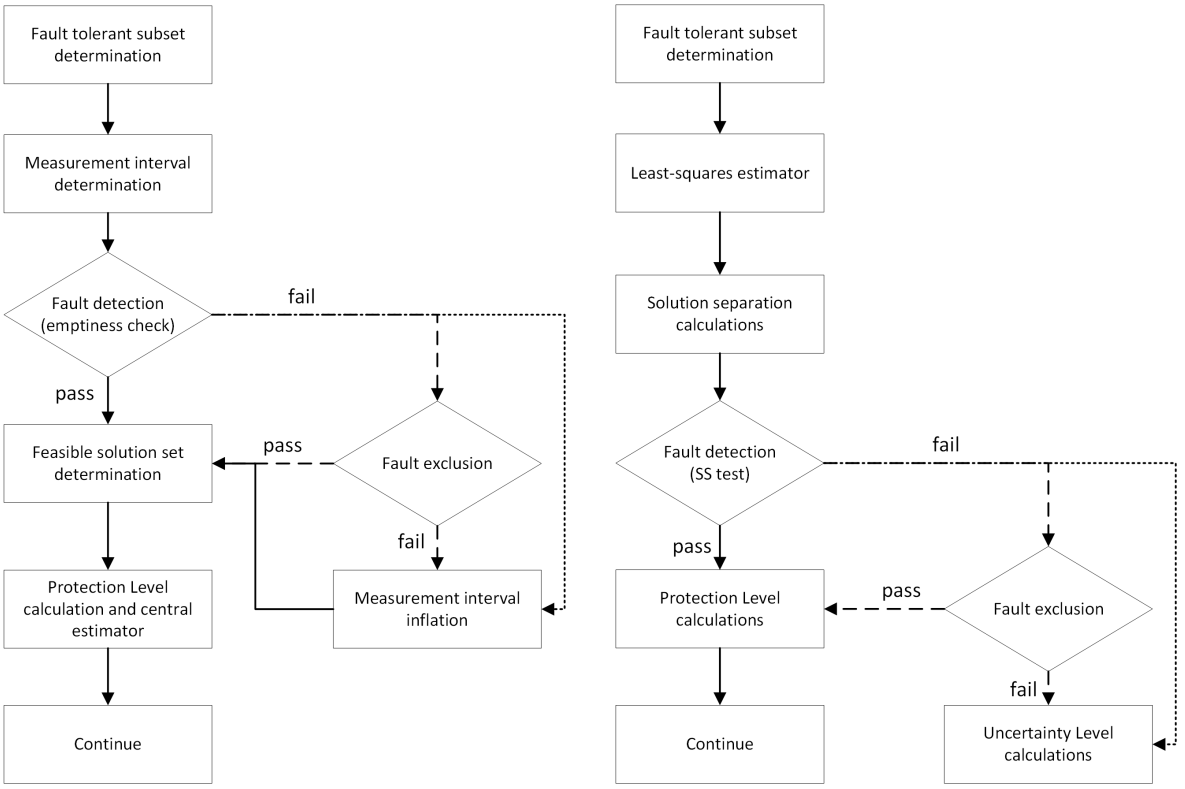


Figure 5.11: Conceptual sketch of the optimal point estimator and associated PL for a two-dimensional example.



(a) Workflow of interval-based FDE

(b) Workflow of ARAIM FDE, adapted from She et al. (2023)

Figure 5.12: Flowcharts of ARAIM and interval-based FDE. In both diagrams, solid arrows apply to either FD or FE purposes, while dotted arrows apply to FD-only and dashed arrows to the entire FDE.

5.2.3 Architecture of the interval-based integrity monitor

To summarize, the workflow of the developed real-time interval-based integrity monitor is given in Fig. 5.12(a). The procedures are introduced as follows:

1. Fault tolerant subset determination. Fault modes with more than r multiple simultaneous faults will not be monitored:
   - Input: P_{sat}, P_{sat_thres}.

- Output: r, $P(H_i)$, P_{NM}.

2. Measurement interval determination:
   a. Compute bounding probabilities P_{Δ_s}, cf. Eq. 5.57.
      - Input: n, r, P_{sat}, P_{NM}, I_{REQ}.
      - Output: P_{Δ_s}.
   b. Determine interval bounds $[\boldsymbol{\Delta}_s]$, cf. Tab. 4.4 and Chap. 3.
      - Input: P_{Δ_s}, $\mathbf{F}$, $[\mathbf{d}]$.
      - Output: $[\boldsymbol{\Delta}_s]$.
3. Set-based fault detection and exclusion (FDE) functions:
   a. Compute the detection polytope $\mathcal{P}_d = \mathcal{P}_{s,0}$ and perform a global emptiness-check (FD function).
   b. If the emptiness check returns a negative result, jump into the next step.
   c. If the emptiness check returns a positive result, activate FE functions by exhaustive local emptiness check against each monitored fault mode.
   d. If any fault is detected but not identified, inflate P_{Δ_s} and determine the new measurement intervals until the global emptiness check can pass.
   e. If any fault is identified, exclude the corresponding measurement from $\mathbf{y}$ and perform Steps 2 to 3 again.

   The output varies with the detection and exclusion results:
   - Input: $\mathbf{A}$, $\mathbf{y}$, $[\boldsymbol{\Delta}_s]$ and possibly r (for FE function).
   - Output:
     - (Step b) None.
     - (Step d) P_{Δ_s} and $[\boldsymbol{\Delta}_s]$.
     - (Step e) $\mathbf{A}$, $\mathbf{y}$, r, P_{Δ_s} and $[\boldsymbol{\Delta}_s]$.
4. Feasible solution set ($\mathcal{S}$) determination, cf. Eq. 5.44.
   - Input: $\mathbf{A}$, $\mathbf{y}$, $[\boldsymbol{\Delta}_s]$, r.
   - Output: $\mathcal{S}$.
5. PL calculation and central estimator, cf. Eq. 5.62.
   - Input: $\mathcal{S}$.
   - Output: $\hat{\mathbf{x}}$, PL.
6. Reference the computed PL to AL and, based on which, notify the user about the service availability.
   - Input: PL, AL.
   - Output: service availability.

The above workflow demonstrates a baseline algorithm for interval-based integrity monitoring. In comparison with the ARAIM FDE process summarized in Fig. 5.12(b), the functions are implemented in a different order, indicating their fundamental difference. The least-squares estimator is implemented initially in ARAIM, and its residuals are used in the FDE functions. In contrast, the interval-based approach employs the central estimator that is derived from its FDE results.

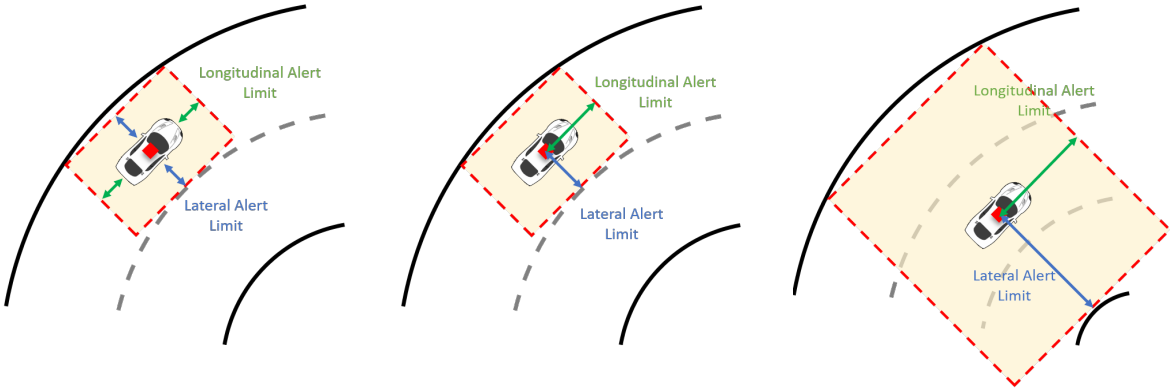
5.2.4 Practical solution to road application

For road transportation applications, the horizontal position is of primary interest. Researchers have developed Alert Limit (AL) concepts for autonomous vehicles in the scenario of US road transportation (Reid et al., 2019, 2023). This section proposes a sub-optimal but practical point estimator compatible with the referenced AL framework for road applications.

Referenced AL framework. The referenced AL concept is based on a bounding box around the vehicle on the road, determined by (i) road width and curvature and (ii) the desired level of situational awareness. Situational awareness in this context is broken down into three categories: (i) which road, (ii) which lane, and (iii) where in the lane, and, hence, can be characterized in three levels:

1. Road determination,
2. Lane determination,
3. In-lane positioning,

which demand various sizes of bounding boxes, cf. Fig. 5.13, which, by definition, can be described as (two-dimensional) interval boxes.



(a) Lane keeping (b) Lane determination (c) Road determination

Figure 5.13: Bounding box geometry for different levels of situation awareness (Reid et al., 2023). Road determination is shown for a 3-lane freeway as an example.

PL for road geometry. A lateral and a longitudinal PLs are expected accordingly. To achieve this, the FSS is projected onto the along-track and cross-track directions, resulting in two circumscribed intervals, i.e., $[\mathcal{S}_{\text{Lon}}]$ and $[\mathcal{S}_{\text{Lat}}]$, respectively:

$$\left[\mathcal{S}_{\text{Lat/Lon}}\right] \coloneqq \bigcup_{i=0}^{h} \left[\mathcal{P}_{\text{Lat/Lon},i}\right], \text{ with } \left[\mathcal{P}_{\text{Lat/Lon},i}\right] = \left\{\mathbf{p}_{\text{Lat/Lon}}^{T} \cdot \mathbf{x} \mid \mathbf{x} \in \mathcal{P}_{s,i}\right\}, \tag{5.63}$$

where $\mathbf{p}_{\text{Lat/Lon}}$ is the projection vector that depends on the road's geometry. For instance, for GPS-only solutions (four unknowns) in an n-frame:

- $\mathbf{p}_{\text{Lat}} = [-\sin\phi,\ \cos\phi,\ 0,\ 0]^T$ for the lateral component,
- $\mathbf{p}_{\text{Lon}} = [\cos\phi,\ \sin\phi,\ 0,\ 0]^T$ for the longitudinal component,

where ϕ denotes the road's orientation, specifically the azimuth angle of the tangent to the road arc. This information is not estimable by the designed algorithm and, hence, should be provided by external sources, e.g., inertial sensors and the steering control system.

The estimated coordinates and PLs can be obtained by following these steps, which substitute Steps 4 to 5 in Sec. 5.2.3:

1. Determine h fault-tolerant polytope solution sets $\mathcal{P}_{s,i}$ for $i = 0, 1, ..., h$;
2. Project $\mathcal{P}_{s,i}$ onto lateral/longitudinal directions, yielding $h+1$ intervals:

$$[\mathcal{P}_{\text{Lat/Lon},i}] \coloneqq [\underline{p}_{\text{Lat/Lon},i}, \overline{p}_{\text{Lat/Lon},i}].$$

3. Determine the interval bounding box's lateral and longitudinal edges:

$$\begin{aligned} [\mathcal{S}_{\text{Lat}}] &\coloneqq [\min(\underline{p}_{\text{Lat},i}), \max(\overline{p}_{\text{Lat},i})], \\ [\mathcal{S}_{\text{Lon}}] &\coloneqq [\min(\underline{p}_{\text{Lon},i}), \max(\overline{p}_{\text{Lon},i})]. \end{aligned} \tag{5.64}$$

4. Employ the central estimator and calculate lateral/longitudinal PLs:

$$\begin{aligned} \hat{\mathbf{x}}_{\mathcal{S},\text{Lat/Lon}} &=\mid p([\mathcal{S}_{\text{Lat/Lon}}]) = \frac{1}{2}\left(\min(\underline{p}_{\text{Lat/Lon},i}) + \max(\overline{p}_{\text{Lat/Lon},i})\right), \\ \text{PL}_{\text{Lat/Lon}} &= \text{rad}([\mathcal{S}_{\text{Lat/Lon}}]) = \frac{1}{2}\left(\max(\overline{p}_{\text{Lat/Lon},i}) - \min(\underline{p}_{\text{Lat/Lon},i})\right). \end{aligned} \tag{5.65}$$

Subsequently, by comparing the computed lateral and longitudinal PLs to corresponding ALs for different levels of situational awareness, the user will be informed about the system availability.

Sub-optimal point estimator. Fig. 5.14 conceptualizes the bounding box derived from the FSS, which are used to determine the point solution (red triangle) and PLs.

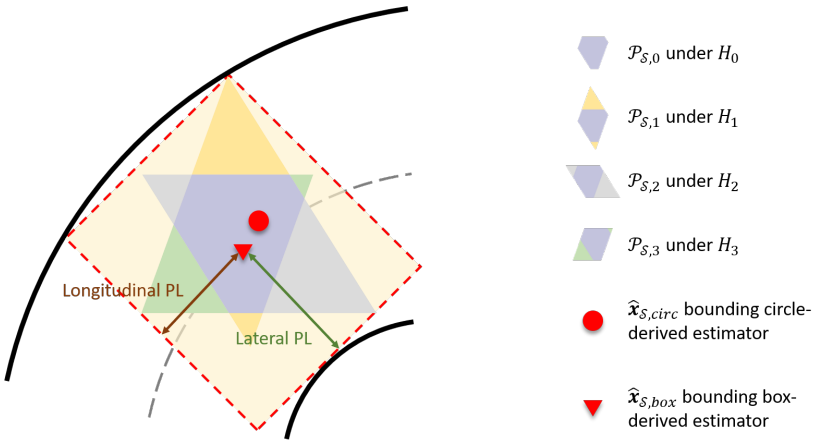


Figure 5.14: Interval box as lateral and longitudinal PLs for vehicles on the road.

The point estimate differs slightly from the solution from the bounding sphere's center (red dot). This is because the latter is optimal in the sense of minimum horizontal positioning error and is unique and invariant under rotation of the coordinate system. In contrast, the bounding box is an arbitrary enclosure of $\mathcal{S}$ and varies with the choice of the orientation of the coordinate system. Hence, the changes in road geometry potentially affect both its shape and scalar measure. In this regard, the point estimator from the bounding box is sub-optimal compared to the one directly from FSS. Schön and Kutterer (2005b) discusses in detail the use and limits of interval boxes as point uncertainty measures for geodetic applications. Despite this, it is advantageous in computational efficiency and fits the referenced AL framework for road applications.

5.3 Performance evaluation and validation

The integrity monitoring algorithms are designed to meet users' requirements concerning both IR and position uncertainty; therefore, it is essential to evaluate both the IR bound and position bound (indicated by PL). The IR bound is compared to a specified IR requirement, while the PL is referenced to AL to assess system availability, which is the percentage of time the navigation solution can be used safely (Joerger and Pervan, 2014). Therefore, deriving a tight IR bound and generating a small PL is critical for ensuring system availability.

This section focuses on evaluating the performance of the developed interval-based integrity monitoring approach in terms of the tightness of its IR bounds. The assessment utilizes both analytical methods and Monte Carlo simulations, comparing the results to classical RAIM approaches through benchmark examples and realistic scenarios.

5.3.1 Analytical evaluation for integrity risk bound: Benchmark problem

The benchmark example (cf. Sec. 5.1.3) is analyzed to compare the interval approach with two RAIM approaches in terms of IR bounds.

Configuration of interval-based approach

A formal description of the FSS $\mathcal{S}$ is fundamental to configuring the interval-based approach. In the benchmark problem, the fault-tolerant solution sets can be expressed explicitly for the null hypothesis (H_0) and each single-measurement fault hypothesis (H_i with $i = 1, 2, 3$):

$$\begin{aligned}
\mathcal{P}_{s,0} &= [\max(y_1, y_2, y_3) - \Delta_s, \min(y_1, y_2, y_3) + \Delta_s], \\
\mathcal{P}_{s,1} &= [\max(y_2, y_3) - \Delta_s, \min(y_2, y_3) + \Delta_s], \\
\mathcal{P}_{s,2} &= [\max(y_1, y_3) - \Delta_s, \min(y_1, y_3) + \Delta_s], \\
\mathcal{P}_{s,3} &= [\max(y_1, y_2) - \Delta_s, \min(y_1, y_2) + \Delta_s],
\end{aligned} \tag{5.66}$$

where the measurement intervals $[-\Delta_s, \Delta_s]$ are identical for three measurements as they all follow normal distributions, and hence, can be determined from the CDF of a normal distribution, cf. Eq. 4.26 to Eq. 4.28.

While $\mathcal{S}$ is determined as the union of the four intervals in Eq. 5.66, the scalar measure $g(\mathcal{S})$ corresponds to the radius of the minimum enclosing interval of $\mathcal{S}$, and the point solution is

represented by its midpoint. Utilizing order statistics leads to the formulation as follows:

$$\begin{aligned} [\mathcal{S}] &= \left[Y_{(2:3)} - \Delta_s, Y_{(2:3)} + \Delta_s\right], \\ g(\mathcal{S}) &\coloneqq \mathrm{rad}([\mathcal{S}]) = \Delta_s, \\ \hat{\mathbf{x}} &\coloneqq \mathrm{cen}([\mathcal{S}]) = Y_{(2:3)}, \end{aligned} \tag{5.67}$$

where $[\mathcal{S}]$ denotes the minimum enclosing interval of $\mathcal{S}$, and $Y_{(2:3)}$ stands for the second order statistic out of three samples. A complete guidance of deriving $[\mathcal{S}]$ is given in Annex. A.1.

Eq. 5.67 suggests that the designed point estimator degenerates to a *median* estimator, with an estimation error bound of Δ_s. This is fundamentally different from its counterpart in the RAIM approaches, which degenerates to a *mean* estimator.

Thereafter, the corresponding IR bound can be computed using Eq. 5.54. Alternatively, the focus can shift to seeking the worst-case fault (WCF) in the sense that the associated HMI probability is maximized, similar to RB RAIM cf. Sec. 2.4.5. For each alternative hypothesis H_i, a worst-case fault vector is defined as $\mathbf{f}_{i,\mathrm{worst}}$, under which the conditional HMI probability given H_i is maximized.

Subsequently, the overall HMI probability can be tightly upper bounded by

$$P(\mathrm{HMI}) \leq P(\mathrm{HMI} \mid H_0) \cdot P(H_0) + \sum_{i=1}^{3} P(\mathrm{HMI} \mid \mathbf{f}_{i,\mathrm{worst}}) \cdot P(H_i). \tag{5.68}$$

The detailed procedure of determining each term in Eq. 5.68 is introduced in Annex. A.2.4.

In the following analysis, the former IR bound is named *NP bound*, while the latter is termed *WCF bound* for differentiation.

It is essential to understand the fundamental difference between the two bounds. The NP bound is established through several inequalities, which can introduce looseness of bounding. In contrast, the WCF bound is directly assessed in relation to the potential worst-case fault, making it a reflection of *realistic* IR bounds. This distinction is further supported by analytical evaluations.

Notably, in the presence of more than three measurements for the scalar problem, the designed point estimator in the interval-based approach is a general *L-estimator*, revealing its inherent robustness. The expected L-estimator is configured by the coefficients r (maximum monitored simultaneous faults) and n (total number of measurements):

$$[\mathcal{S}] = \left[Y_{(n-r:n)} - \Delta_s, Y_{(r+1:n)} + \Delta_s\right], \ \hat{\mathbf{x}} \coloneqq \mathrm{cen}([\mathcal{S}]) = \frac{1}{2} \cdot \left(Y_{(n-r:n)} + Y_{(r+1:n)}\right). \tag{5.69}$$

Configuration of RAIM approaches

The derivations in Joerger et al. (2014) serve as a foundation for the following analysis:

- For SS RAIM, the FA probability requirement is equally allocated among the three test statistics, which is deemed optimal in this case. The IR bound is then determined using Eq. 2.93.
- For RB RAIM, the worst-case fault magnitude is computed following the instruction provided (using a line search method). Subsequently, the IR bound can be determined based on Eq. 2.67.

Relevant parameters are summarized in Tab. 5.2 and utilized in the evaluation. Notably, the parameter Δ_d (for set-based detection) is exclusively for evaluation purposes but not necessarily needed in the integrity monitoring algorithm.

Comparative analysis

Fig. 5.15 compares four IR bounds for the three approaches:

- NP bound for the interval approach (blue-colored curves),
- WCF bound for the interval approach (green-colored curves),
- IR bound for RB RAIM (red-colored curves),
- IR bound for SS RAIM (black-colored curves).

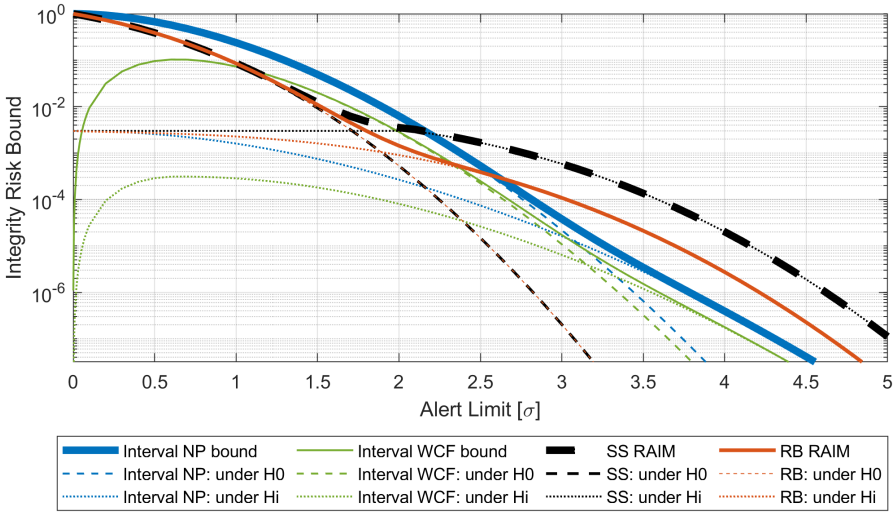


Figure 5.15: Integrity risk bound comparison among the interval based approach, SS RAIM and RB RAIM for the benchmark example.

The IR bounds are displayed with respect to various AL values, allowing for performance comparison in terms of bounding tightness. Specifically, lower IR bound values indicate tighter bounds for a given AL. Such tight bounds are desirable, as they contribute to enhanced system availability.

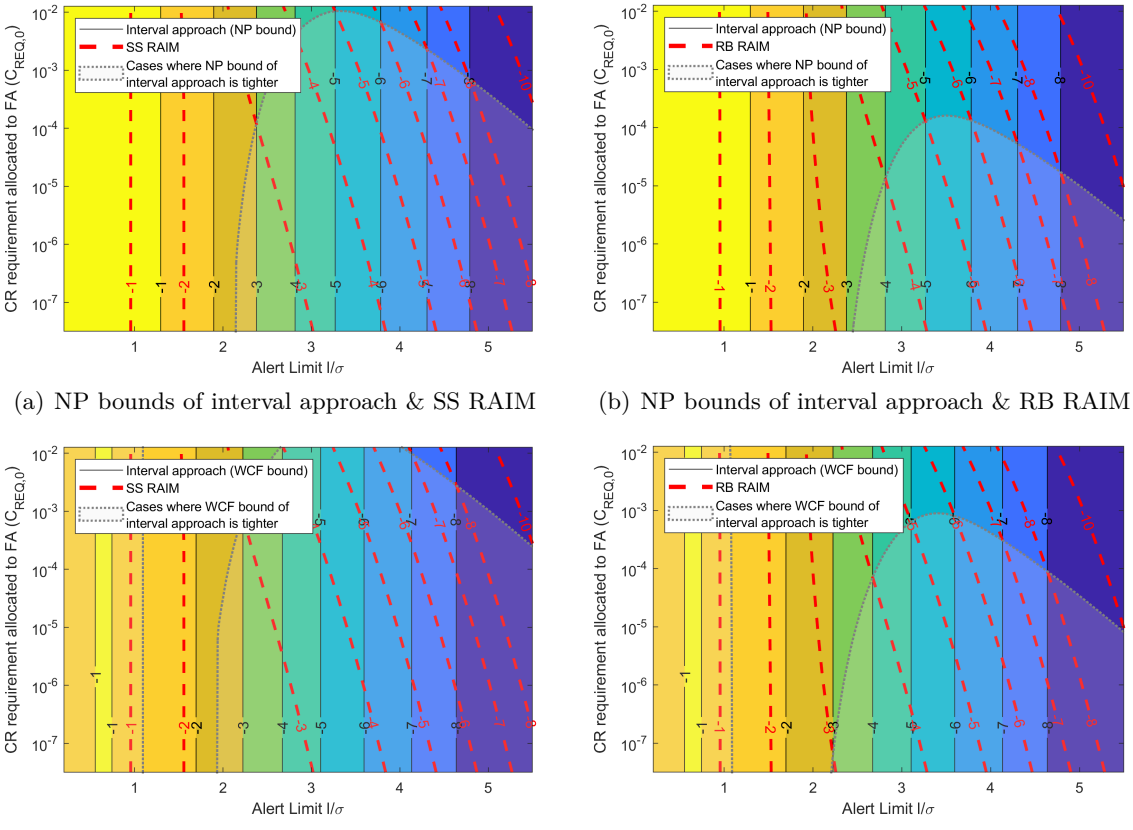
Comparing the NP and WCF bounds for the interval approach, there is no doubt that the WCF bound is tighter. The main reason is discussed in Sec. 5.2.1 that two inequalities for the NP bound may introduce conservatism, i.e., the ignorance of the impact of ND event (Eq. 5.32) and using NP probability to bound PF probability (Eq. 5.37). The difference is most significant when AL approaches zero. As shown in Fig. 5.15, the blue curve representing the NP bound starts from one while the green curve for the WCF bound from zero. This is due to the impact of ND event, which is virtually not happening at AL $= 0$. The WCF bound gets closer to the NP bound as AL grows. The peak value of the WCF bound appears at AL $\approx 0.6\,\sigma$. Both bounds approach zero at infinity.

Additionally, it can be noticed that the two RAIM approaches have tighter bounds than the WCF bound of interval approach when the AL is only between around $1.1\,\sigma$ to $1.9\,\sigma$ (RB) or to $2.4\,\sigma$ (SS), while interval-based approach outperforms for smaller and larger ALs. The NP bound of the interval approach is less but still advantageous. For instance, it outperforms the two RAIM approaches when the AL is greater than $2.2\,\sigma$ (SS) and $2.6\,\sigma$ (RB). In this benchmark example, at larger AL values, the IR bounds for the interval approach are up to two orders of magnitude tighter than the SS bound and one order than the RB bound.

To better understand the different behaviors, the bounds based on Eq. 5.39, Eq. 5.68, Eq. 2.93 and Eq. 2.67 are broken down into two components for each approach, i.e. the fault-free hypothesis H_0 and the sum of probabilities over all faulty hypotheses H_i, $i = 1, 2, 3$.

The former is displayed as dashed curves, dominating the total IR bound for smaller AL values. Meanwhile, the latter has a greater impact on the total budget for larger AL values in all approaches. The major difference lies in the criterion, i.e., the switch point of dominant component, is greater for the interval-based approach than the SS RAIM and RB RAIM, i.e., around $3.2\,\sigma$ (WCF bound), $3.1\,\sigma$ (NP bound), $1.7\,\sigma$ (SS) and $1.9\,\sigma$ (RB).

The sensitivity of the IR bounds for RB RAIM, SS RAIM and interval approach is further investigated in Fig. 5.16, where the AL and $C_{REQ,0}$ (the CR requirement allocated to FA) are varied for the benchmark example. The two bounds of the interval approach are independent of $C_{REQ,0}$. Both outperform the RB and SS bounds for stricter $C_{REQ,0}$ and/or greater AL (ℓ/σ), i.e, in the lower-right portion of Fig. 5.16(a) to Fig. 5.16(d). This trend is valuable because those cases where the interval-based approach can provide tighter IR bounds are usually of particular interest to integrity applications. Moreover, the advantage of WCF bounds of the interval approach at lower ALs remains for all $C_{REQ,0}$, which is visible in Fig. 5.16(c) and Fig. 5.16(d).



(a) NP bounds of interval approach & SS RAIM

(b) NP bounds of interval approach & RB RAIM

(c) WCF bounds of interval approach & SS RAIM

(d) WCF bounds of interval approach & RB RAIM

Figure 5.16: Contour plots: IR bounds with respect to the Alert Limit and continuity risk requirements for the interval approach and RAIM approaches for the benchmark example.

5.3.2 Analytical evaluation for integrity risk bound: Hybrid error models

In addition to purely Gaussian-distributed errors, further hybrid error models should be evaluated, accounting for more complicated scenarios. The following scenarios are analyzed:

- **Scenario 1:** The observation errors consist of two components $\mathbf{e} = \mathbf{e}_1 + \mathbf{e}_2$. One component follows a normal distribution, i.e., $\mathbf{e}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The other component $\mathbf{e}_2$ is a deterministic bias that is known to be upper bounded by an interval $[-1, 1]$.
- **Scenario 2:** The observation errors follow identically and independently long-tailed distributions. This evaluation uses a student's t-distributed error model $\mathbf{e} \sim t(\nu = 10)$.

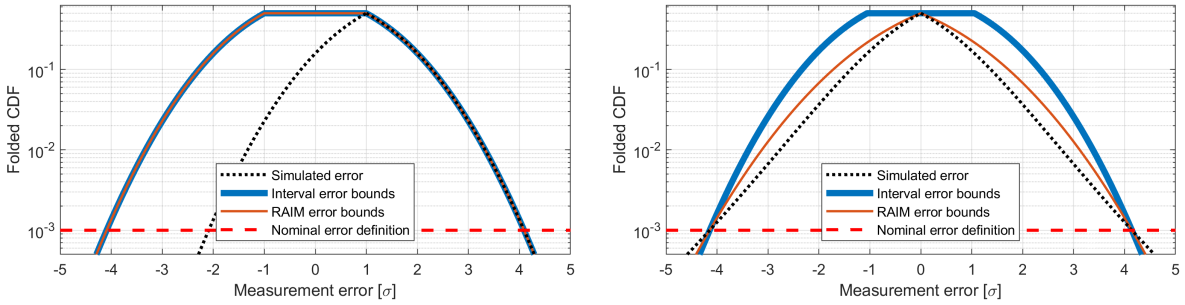
The evaluation method is the same as in Sec. 5.3.1, comparing the NP bound of the interval approach with the IR bounds of RB and SS RAIM.

Following the ARAIM nominal error concepts (cf. Sec. 2.4.6), the measurement error bounds for RB/SS RAIM are configrued with two parameters σ_{ob} (zero-mean overbounding Gaussian standard deviation) and b_{nom} (worst-case nominal bias). To construct measurement intervals for the interval-based approach, two parameters are defined, i.e., σ_{ub} (standard deviation describing random errors) and b_{ub} (interval radius describing remaining systematic errors). This leads to the radius $\Delta_s = k_\Delta\, \sigma_{ub} + b_{ub}$. Tab. 5.3 shows the specified values for these parameters, aimed at achieving CDF overbounding based on the principles described in Sec. 2.4.2.

Table 5.3: Overbounding parameters used in different approaches for IR bounds evaluation (units: σ).

	Scenario 1	**Scenario 2**
RB/SS RAIM	$\sigma_{ob} = 1, b_{nom} = 1$	$\sigma_{ob} = 1.341, b_{nom} = 0$
Interval-based approach	$\sigma_{ub} = 1, b_{ub} = 1$	$\sigma_{ub} = 1, b_{ub} = 1.054$

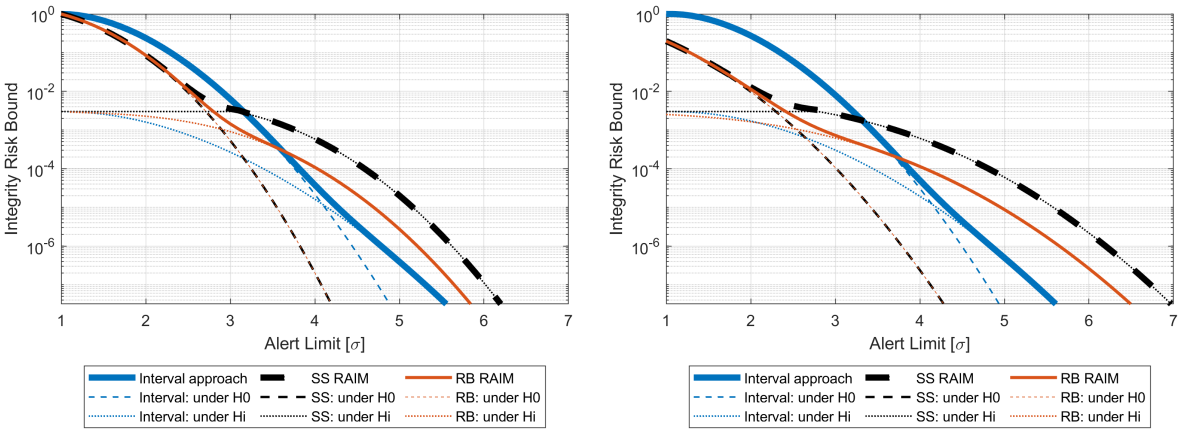
Fig. 5.17 presents the folded CDF for hybrid error models along with established error bounds. Notably, the illustration principles are consistent with those in Fig. 4.6, with the key difference being the probability that errors are guaranteed to be bounded. While Fig. 4.6 aims to bound all simulated errors for a fault-free analysis, Fig. 5.17 restricts its probabilistic bounds down to $P_{sat} = 10^{-3}$.



(a) Scenario 1 (b) Scenario 2

Figure 5.17: Folded CDF for two scenarios of hybrid error models and associated bounds.

In Scenario 1, two approaches share identical error bounds, and consequently, in Fig. 5.18(a), they demonstrate unchanged tightness of bounding (compared to Fig. 5.15) with enlarged ALs. In contrast, different error bounding strategies are applied in Scenario 2, i.e., the remaining systematic error that results in the long tails is handled deterministically in the interval-based approach, while both strategies bound the tail distribution of simulated errors down to 10^{-3}. As a consequence, the bounding tightness differs from Scenario 1, i.e., both the disadvantage of the interval approach at smaller ALs (or equivalently, higher IR) and its advantage at greater ALs (lower IR) appear more significant cf. Fig. 5.18(b). In conclusion, the interval bounding for remaining systematic errors benefits the interval-based integrity monitoring approach, contributing to tighter IR bounds under certain conditions, e.g., critical IR requirements.



(a) Scenario 1 (b) Scenario 2

Figure 5.18: IR bound comparison among the interval based approach, SS RAIM and RB RAIM for two scenarios of hybrid error models.

5.3.3 Monte Carlo evaluation for realistic scenarios

In order to understand whether the difference between the interval-based and RAIM approaches is still significant in more realistic scenarios, an example with joint GPS and Galileo systems in Blanch et al. (2015a) and WG-C ARAIM TSG (2019) is assessed in the following analysis. Besides, the IR bounds are evaluated over 24 hours at an example Hannover location. Comparative analysis is conducted for both scenarios.

Benchmark dual-constellation example

Consider the following geometry matrix $\mathbf{A}$ (for 5 GPS observations followed by 5 Galileo observations) and the variances of measurement errors contained in Σ:

$$\mathbf{A} = \begin{bmatrix} 0.023 & 0.995 & -0.097 & 1 & 0 \\ 0.675 & -0.690 & -0.261 & 1 & 0 \\ 0.072 & -0.660 & -0.748 & 1 & 0 \\ -0.940 & 0.255 & -0.227 & 1 & 0 \\ 0.591 & -0.754 & -0.288 & 1 & 0 \\ -0.324 & -0.035 & -0.946 & 0 & 1 \\ -0.675 & 0.436 & -0.596 & 0 & 1 \\ 0.094 & -0.700 & -0.708 & 0 & 1 \\ 0.557 & 0.309 & -0.771 & 0 & 1 \\ 0.662 & 0.696 & -0.278 & 0 & 1 \end{bmatrix}, \quad \Sigma = \mathrm{diag}\left(\begin{bmatrix} 3.887 \\ 1.438 \\ 0.860 \\ 1.638 \\ 1.323 \\ 0.843 \\ 0.896 \\ 0.867 \\ 0.857 \\ 1.362 \end{bmatrix} \right). \tag{5.70}$$

A skyplot is shown in Fig. 5.19 for five GPS satellites (denoted by G1 to G5) and five Galileo satellites (E6 to E10). In addition to the full set of ten satellites, five subsets are chosen for analysis, with corresponding excluded satellites highlighted.

A total of 100 Monte Carlo trials were conducted, each under six scenarios, including a full set and various subsets. Observation values were generated randomly based on Gaussian distributions, with variances in Eq. 5.70. It is noteworthy that faulty measurements were not particularly considered in simulations for the following reasons: (i) the satellite failure probability was set $P_{sat} = 10^{-5}$, which is fairly low given the number of Monte Carlo runs; (ii) fault detection is not the focus of this study.

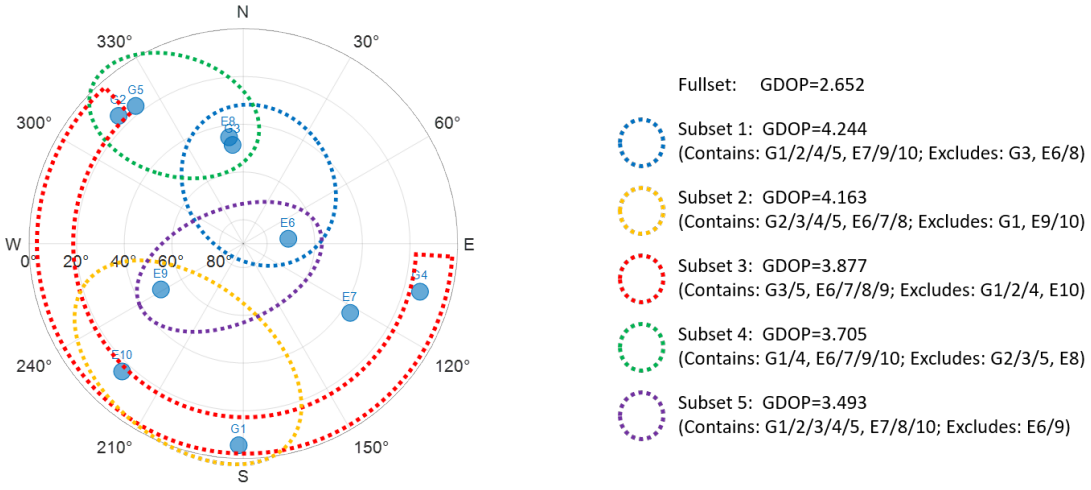


Figure 5.19: Skyplot of the benchmark dual-constellation example. The full set and five subsets out of the ten satellites are used in the comparative analysis. The excluded satellites in each scenario are highlighted.

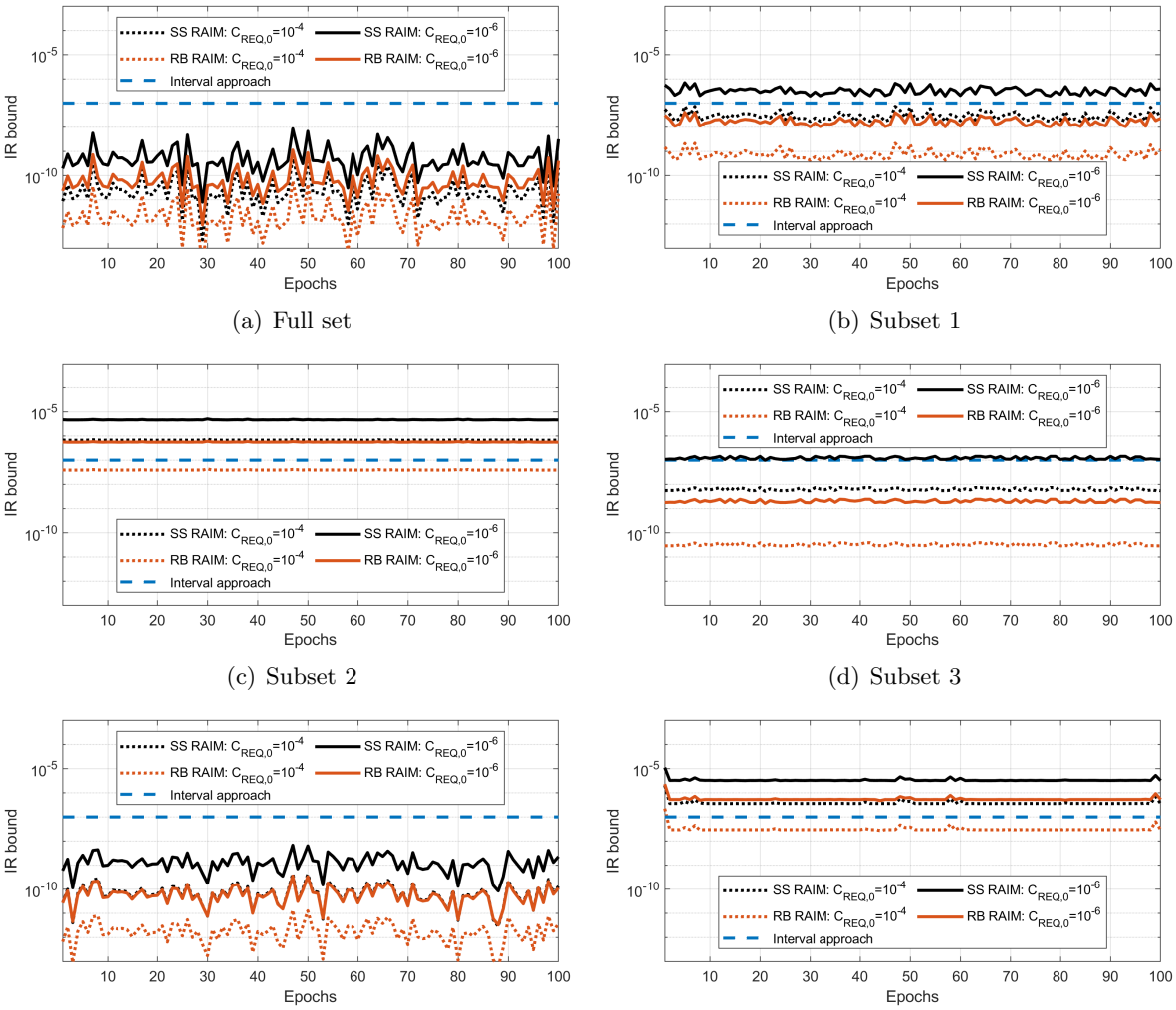
With the IR requirement set to 10^{-7}, the FSS of the interval-based approach can be computed using the method in Sec. 5.2.2. To establish a baseline for comparison, state-wise PLs were calculated. Subsequently, the IR bounds for RB and SS RAIM can be evaluated with respect to these PLs following the same method as outlined in Sec. 5.3.1, assuming only single-measurement faults and no risk due to not-monitored fault modes ($P_{NM} = 0$). The final IR is the sum of IR from all states. In this context, the three approaches are aligned in the position domain, corresponding to the same PL or AL, based on the principles used to evaluate scalar estimation problems discussed in Sec. 5.3.1 and Sec. 5.3.2. A tight IR bound is consistently preferred.

This strategy is free of dependence on PL computation during evaluation, thereby avoiding additional variability in performance results that may arise from differences in how RAIM methods implement PL calculations.

Fig. 5.20 displays the comparison. The RB and SS RAIM are evaluated for $C_{REQ,0} = 10^{-4}$ and $C_{REQ,0} = 10^{-6}$. Despite the advantage of RB RAIM over SS RAIM, the interval-based approach outperforms the other two approaches partially, i.e., for $C_{REQ,0} = 10^{-6}$, Subsets 1/3 (SS) and Subsets 2/5 (both), while for $C_{REQ,0} = 10^{-4}$, Subsets 2/5 (SS only). Hence, the following influence factors may be summarized concerning the performance of the interval-based approach:

- **CR requirement.** The interval approach may be advantageous in the case of stringent CR requirements (i.e., low probabilities). This finding is consistent with the analysis for the benchmark scalar problem.
- **Satellite geometry.** The interval approach outperforms the RAIM approaches under certain geometries, which is not necessarily reflected by the DOP values.
- **Satellite numbers (interval width).** Despite wider intervals required due to fewer excluded satellites in Subsets 1/2/5, the interval approach still provides relatively tight IR bounds, suggesting that the satellite number (or equivalently the interval width) does not necessarily dominate the bounding performance.

In conclusion, in a realistic scenario under pure stochastic measurement error bounding, the advantage of the interval-based approach over classical RB/SS RAIM approaches may not be



(a) Full set (b) Subset 1

(c) Subset 2 (d) Subset 3

(e) Subset 4 (f) Subset 5

Figure 5.20: IR bound comparison among the interval based approach, SS RAIM and RB RAIM for a benchmark dual-constellation example. The full set and five subsets out of ten satellites are taken into analysis.

assured significant, which is jointly affected by the satellite geometry, satellite numbers, as well as the CR requirement for RAIM approaches.

24-hours evaluation

For a comprehensive evaluation that takes into account all possible geometries and measurement-level error bounding, an analysis was conducted for an example location in Hannover, Germany, for 24 hours on December 9th, 2021. The assumptions and procedures for analyzing the dual-constellation example apply to this study, except that different measurement-level error bounding approaches are used for the interval-based approach and for RB/SS RAIM. Tab. 5.4 showcases the management of six major error sources, with residual tropospheric and ionospheric errors being treated differently.

Fig. 5.21(a) and Fig. 5.21(b) depict the IR bound comparison when using GPS or GPS/Galileo satellites for single-frequency positioning. Although the RB RAIM outperforms SS RAIM, the difference is relatively tiny at most epochs. Both provide looser IR bounds than the interval approach (set to $I_{REQ} = 10^{-7}$), the dominant driver for which is the utilization of interval bounding for residual ionospheric and tropospheric errors. Using more satellites (from

Table 5.4: Nominal error modeling for RB/SS RAIM and the interval-based approach in the 24-hours IR evaluation. Two positioning modes are investigated: single-frequency (SF) and dual-frequency (DF).

Nominal errors	RB/SS RAIM	Interval approach
SISRE (URA specs.)	GPS in NAV data, Galileo SF: 7.5 m, DF: 6 m	
Residual tropospheric error[1]	$\sigma_{\mathrm{TRP}} = 0.12 \cdot \frac{1.001}{\sqrt{0.002001+\sin^2\theta}}\,\mathrm{m}$	Interval bounds (Sec. 3.2)
Residual ionospheric error[2]	SF: $\sigma^2_{\mathrm{UIRE}} = \max\left\{\left(\frac{I_r^k}{5}\right)^2, (F_{pp}\,\tau_{vert})^2\right\}$ DF: not considered	SF: Interval bounds (Sec. 3.3) DF: not considered
Multipath error	not considered	
Receiver code noise[1]	$\sigma_{\mathrm{REC}} = \frac{\sigma_0}{\sin\theta}$ with σ_0: GPS SF 0.1577 m & DF 0.4246 m, Galileo SF 0.1373 m & DF 0.3371 m	
Nominal signal deformation	SF: not considered, DF: $b_{nom} = 0.75\,\mathrm{m}$	

[1] θ is the satellite elevation angle.
[2] In the MOPS User Ionospheric Range Error (UIRE) model, I_r^k is the ionospheric delay correction in meters, F_{pp} is the obliquity factor and τ_{vert} is a geomagnetic latitude-dependent value; a detailed description is referred to RTCA/DO-229D (2006).

GPS-only to GPS/Galileo) can reduce the difference because of the different behaviors with respect to stochasticity of the interval approach and the least-squares estimator that RB/SS RAIM is built on. Specifying a less strict $C_{REQ,0}$ makes little improvement.

In the dual-frequency cases (Fig. 5.21(c) and Fig. 5.21(d)), especially the dual-constellation case, RB/SS RAIM provide tighter IR bounds than the interval approach due to the elimination of ionospheric delays. The interval approach is yet to be improved, e.g., by further enhancing the measurement-level interval bounding for various error sources.

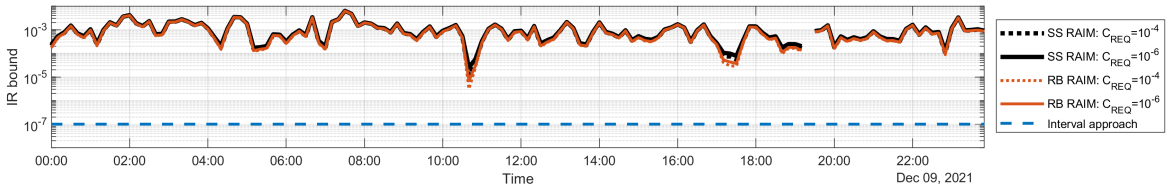
This evaluation relies on certain pre-conditions and assumptions, and hence, the performance of these approaches in real-world applications may differ case by case:

- **Risk allocation.** The CR is allocated equally to all states for SS RAIM; and, for both RB/SS RAIM, the IR budget is contributed by all states for fair comparison. These approaches, by design and in practice, are only concerned with three positional states. Moreover, the IR is usually allocated arbitrarily in prior for specific applications.
- **Multiple hypotheses.** This evaluation considered only single-measurement faults and ignored the probability of not-monitored fault modes for simplicity. Even more complexities can arise in real-world scenarios.
- **Impact of fault detection.** Faulty measurements were not particularly simulated in the Monte Carlo evaluation, which may also have an impact.

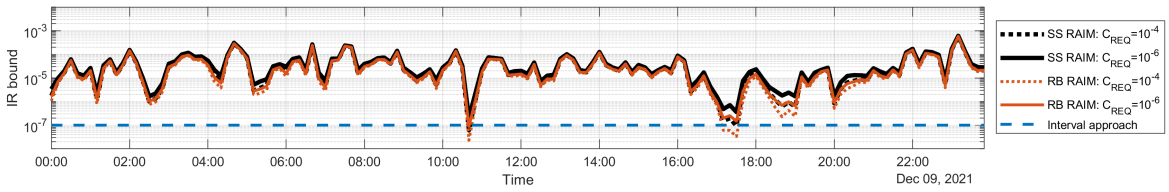
Furthermore, the simulation is based on an ideal satellite availability, which is easily violated in reality, especially in urban environments. Less tracked satellites will significantly degrade the performance of RAIM approaches. The practical significance of the interval approach is further explored in Sec. 6.2.

5.3.4 Discussion

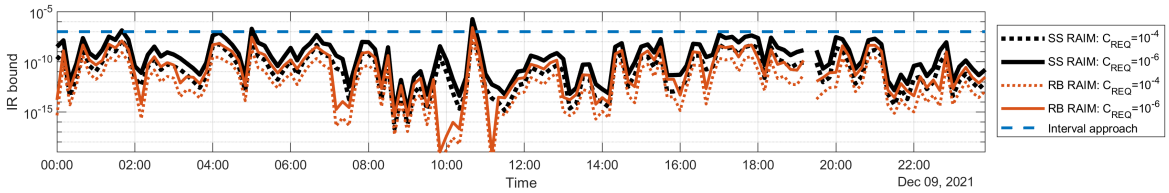
Benefiting from the interval-represented uncertainty modeling for GNSS pseudorange measurements, an alternative autonomous integrity monitoring approach has been designed in



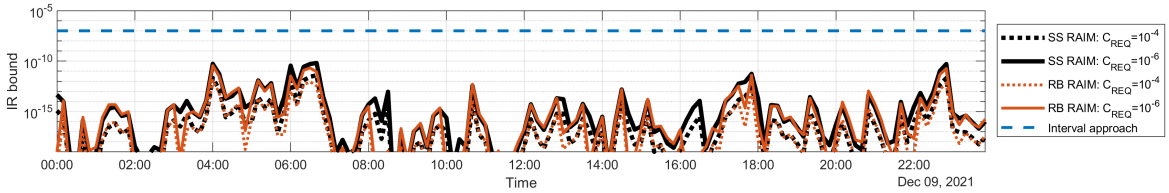
(a) GPS-only, single-frequency



(b) GPS+Galileo, single-frequency



(c) GPS-only, dual-frequency



(d) GPS+Galileo, dual-frequency

Figure 5.21: IR bound comparison among the interval based approach, SS RAIM and RB RAIM for the 24-hours IR evaluation. Two positioning modes are investigated: single-frequency and dual-frequency.

this chapter based on the set-based fault detection and exclusion and set-based central estimator. By incorporating stochastic methods, such as Gaussian overbounding, and deterministic interval bounding methods, the interval-based approach offers a measure of safety that can be interpreted probabilistically for integrity purposes. It adopts the multiple hypotheses framework from ARAIM and, hence, can protect the navigation system against multiple simultaneous faults, which are likely encountered in urban navigation scenarios.

The designed approach is evaluated and compared to classical approaches, including RB/SS RAIM, in terms of tightness of IR bounds through a benchmark scalar problem analytically; besides, Monte Carlo simulations are performed for analyzing more realistic scenarios. Under various error models, the interval approach demonstrates superior performance in certain conditions in terms of tight IR bounding in the benchmark problem. In the Monte Carlo evaluations, it outperforms the classical approaches for single-frequency positioning applications when equipped with the developed measurement-level interval bounds while yet to be improved for dual-frequency scenarios.

6

Application examples and performance analysis

This chapter is dedicated to demonstrating potential applications of the strategies investigated concerning error bounding, uncertainty modeling and propagation, and GNSS integrity monitoring. Specifically, it explores the potential improvements to the baseline ARAIM algorithm and the implementation of real-time integrity monitoring for urban navigation using the interval-based approach.

6.1 Improved fault-free Horizontal Protection Level for aviation users

The error bounding approaches for the least-squares estimator can be applied in ARAIM to calculate the HPL that is intended for the fault-free hypothesis (H_0). This is referred to as fault-free HPL in subsequent discussions. This modification leads to improvements by providing a reduced HPL compared to the baseline ARAIM algorithm. This section evaluates the performance of this approach at both local and global levels.

6.1.1 Methodology

In the baseline ARAIM algorithm, the HPL is computed as the norm of two state-wise PLs, accounting for multiple hypotheses, cf. Sec. 2.4.6. The state-wise fault-free PL consists of two components: the uncertainty due to stochasticity, represented by a confidence interval, and the impact of nominal bias, which is captured by worst-case interval values:

$$HPL_{0,(m)} = k_{0,(m)} \cdot \sigma_{0,(m))} + |\mathbf{k}_{0,(m)}^T|\, \mathbf{b}_{nom}, \text{ with } m = 1, 2, \tag{6.1}$$

$$HPL_0^{BL} = \sqrt{HPL_{0,(1)}^2 + HPL_{0,(2)}^2}, \tag{6.2}$$

where $k_{0,(m)}$ is a coefficient based on the IR requirement allocated to state m, $\sigma_{0,(m)}$ represents the standard deviation of the estimation, $\mathbf{k}_{0,(m)}^T$ and $\mathbf{b}_{nom}$ denote the m-th row vector of the least-squares transformation matrix $\mathbf{K}$ and nominal bias vector, respectively (defined in Eq. 2.36 and Eq. 2.96).

While Eq. 6.1 provides a tight bound within each state, the norm operation in Eq. 6.2 can introduce conservatism in horizontal position bound calculation, for example, overestimating the impact of nominal bias (Pagot et al., 2017; Nikiforov, 2019).

The error bounding approaches for the least-squares estimator, as outlined in Sec. 4.2, offer an alternative scheme for combining uncertainties utilizing set theory and directly operating

on the horizontal plane. This modifies Eq. 6.1 and Eq. 6.2 into

$$HPL_0^{\mathcal{Z}} = \sqrt{\left(k_{0,(1)} \cdot \sigma_{0,(1)}\right)^2 + \left(k_{0,(2)} \cdot \sigma_{0,(2)}\right)^2} + \mathrm{rad}\left(\mathcal{Z}(\hat{\mathbf{x}}, \mathbf{K}\,\mathrm{diag}\left(\mathbf{b}_{nom}\right))\right). \tag{6.3}$$

The first term remains unchanged, calculating the norm for the stochastical component, while the second term, the zonotope's radius, accounts for the uncertainty due to nominal bias.

The subsequent simulation analysis evaluates the reduction by the zonotope-based $HPL_0^{\mathcal{Z}}$ from the baseline HPL_0^{BL}:

$$\delta HPL_0 = HPL_0^{BL} - HPL_0^{\mathcal{Z}}. \tag{6.4}$$

The objective is to assess the impact of nominal bias on HPL computation without altering the treatment of stochasticity in the baseline ARAIM algorithm.

6.1.2 Simulation analysis

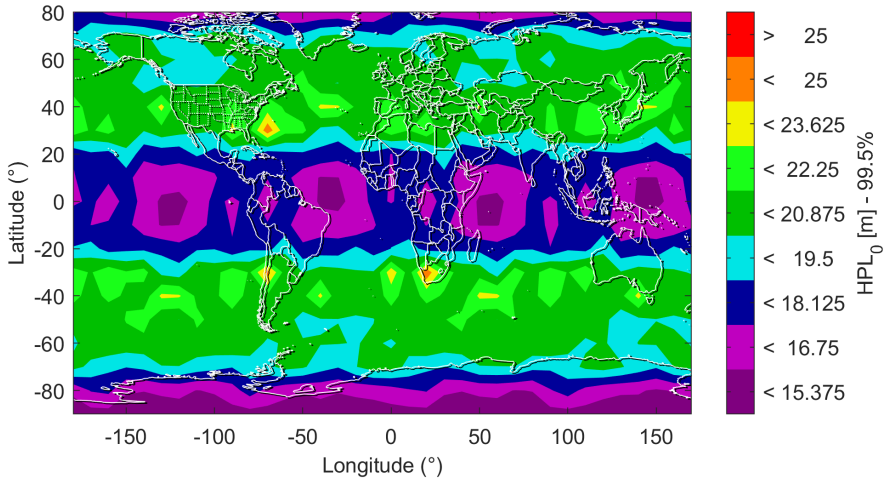
The simulation tool was developed based on the Stanford MatLab Algorithm Availability Simulation Tool (MAAST) for ARAIM Version 0.4 (available at https://gps.stanford.edu/resources/software-tools/maast (Jan et al., 2001)). Tab. 6.1 summarizes the configuration of related parameters in the simulation analysis. Particularly, the horizontal IR requirement is set to 10^{-8}, and a budget of 0.5×10^{-8} is allocated to each of the two states, resulting in $k_{0,(m)} = 5.847$.

Table 6.1: Configuration of related parameters for the fault-free HPL simulation analysis.

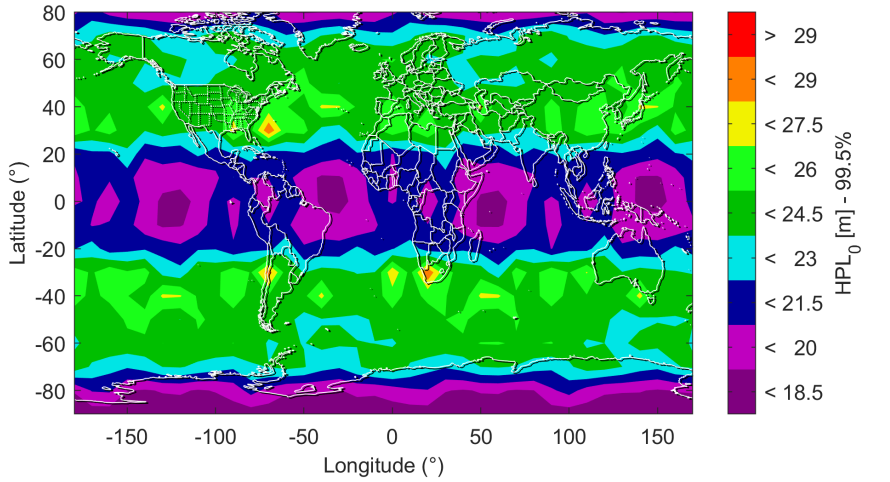
Parameter	Value
Satellite elevation mask	5°
Horizontal integrity risk requirement	10^{-8}
URA for GPS satellites	2.4 m
URA for Galileo satellites	6 m
Nominal bias for GPS satellites (b_{nom})	0.75 m and 2 m
Nominal bias for Galileo satellites (b_{nom})	0.75 m and 2 m
Sampling time interval	300 s
Simulation duration	1 week

Two scenarios were simulated using all available GPS and Galileo satellites for comparison, with nominal biases of $b_{nom} = 0.75\,\mathrm{m}$ and $b_{nom} = 2\,\mathrm{m}$. The results shown in Fig. 6.1, Fig. 6.2 and Fig. 6.3 reveals the following insights:

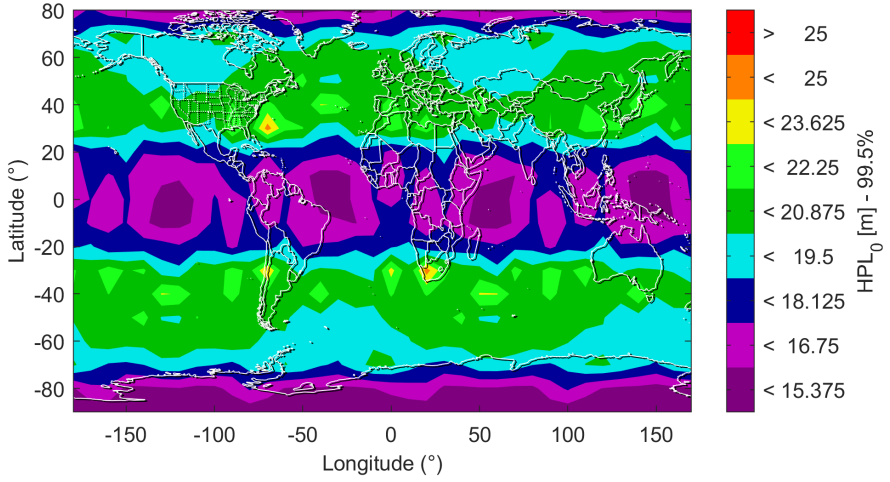
- The zonotope-based method consistently reduces the computed fault-free HPL, both geographically and temporally. The reductions reach up to approximately 0.65 m for $b_{nom} = 0.75\,\mathrm{m}$ and around 1.7 m for $b_{nom} = 2\,\mathrm{m}$.
- The degree of reduction is primarily influenced by the predefined b_{nom} values and the measurement geometry. Greater b_{nom} values lead to more significant improvements, while the global distribution pattern remains relatively consistent across different b_{nom} settings.
- It is inconclusive which regions benefit the most from the zonotope-based HPL computation. Interestingly, both the largest improvement (maximum δHPL_0) and the smallest reduction (minimum δHPL_0) occur in the mid-latitude regions, with the low-latitude and high-latitude regions showing moderate improvements with the zonotope method.



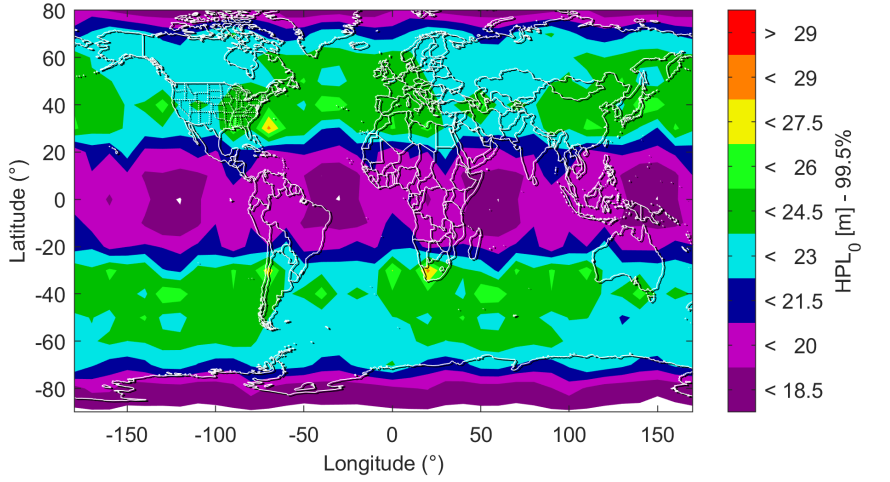
(a) HPL_0^{BL} ($b_{nom} = 0.75\,\mathrm{m}$)



(b) HPL_0^{BL} ($b_{nom} = 2\,\mathrm{m}$)

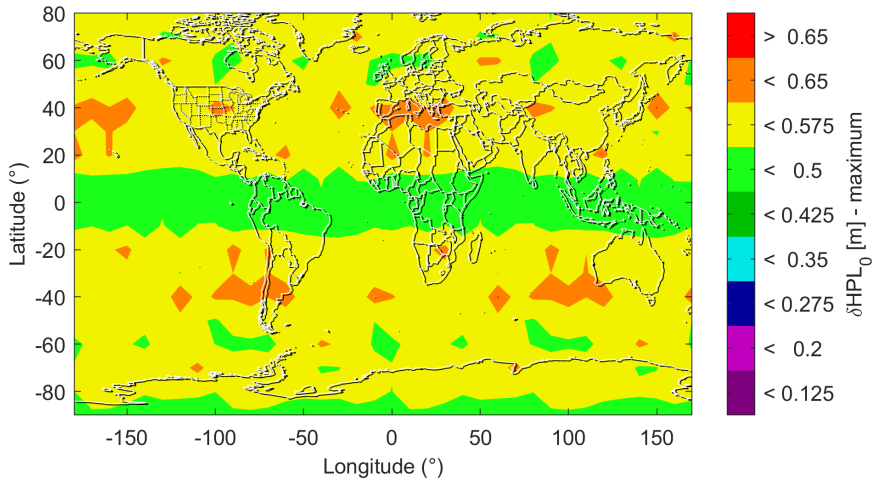


(c) HPL_0^{Z} ($b_{nom} = 0.75\,\mathrm{m}$)

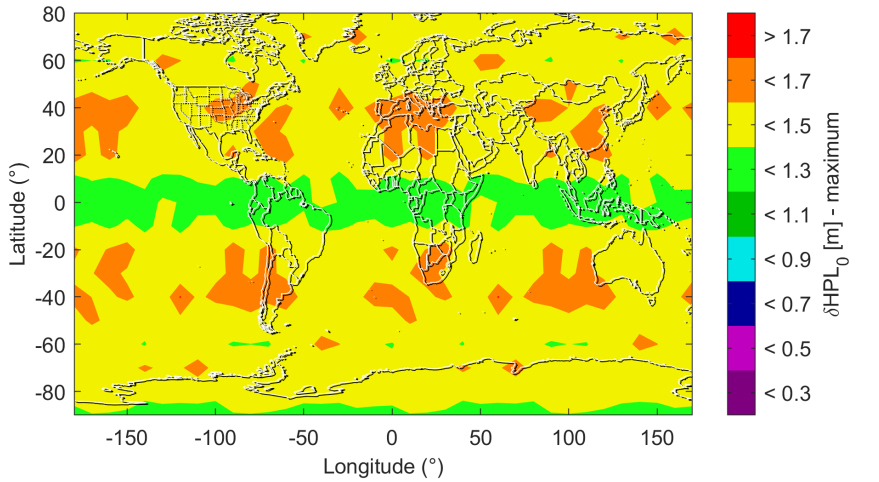


(d) HPL_0^{Z} ($b_{nom} = 2\,\mathrm{m}$)

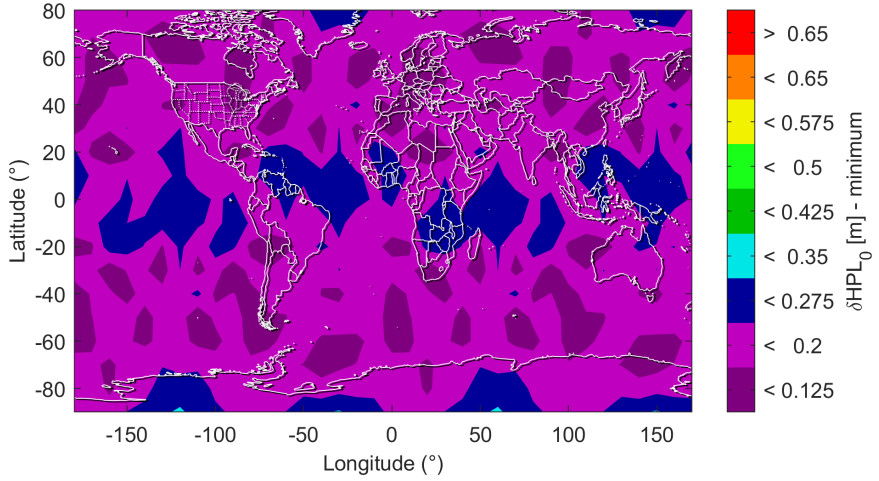
Figure 6.1: Global contours of simulation results for the fault-free HPL using two methods, denoted by HPL_0^{BL} for the baseline ARAIM algorithm and HPL_0^{Z} for the zonotope-based method. Two scenarios concerning the predefined nominal bias are simulated: $b_{nom} = 0.75\,\mathrm{m}$ and $b_{nom} = 2\,\mathrm{m}$.



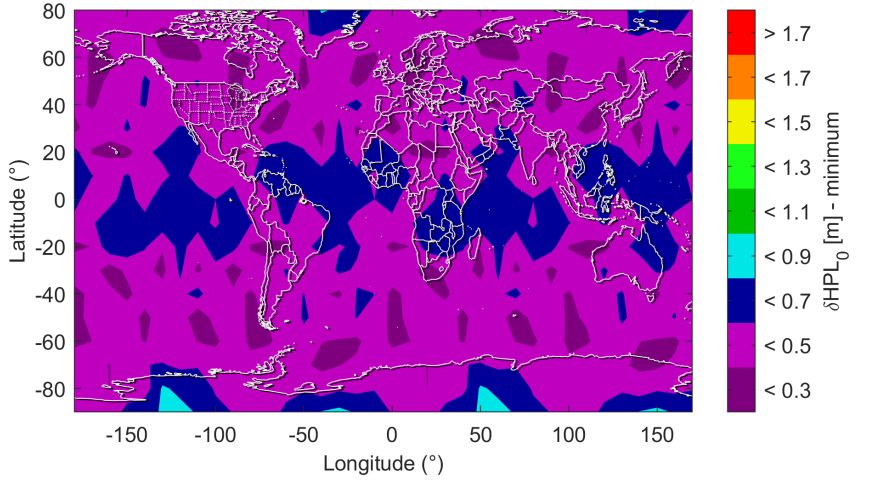
(a) Maximal δHPL_0 ($b_{nom} = 0.75\,\mathrm{m}$)



(b) Maximal δHPL_0 ($b_{nom} = 2\,\mathrm{m}$)

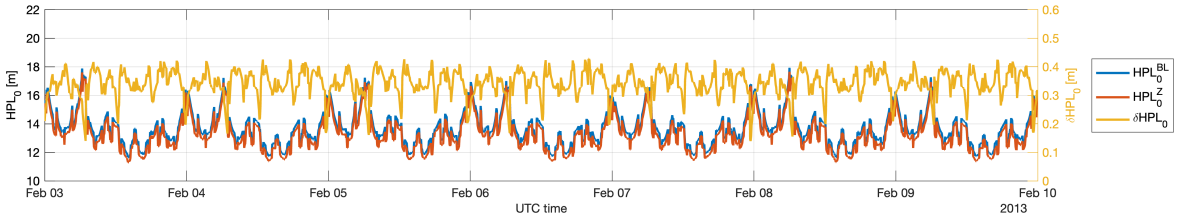


(c) Minimal δHPL_0 ($b_{nom} = 0.75\,\mathrm{m}$)

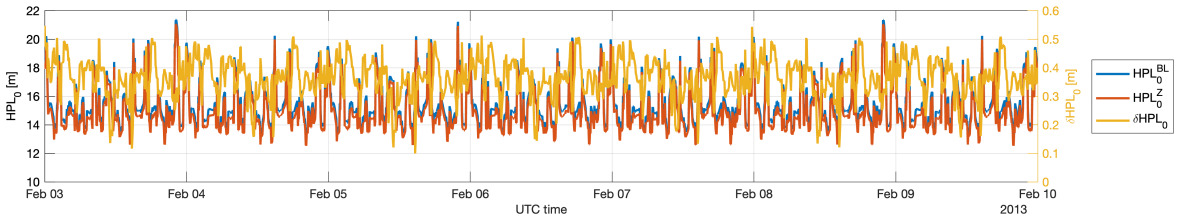


(d) Minimal δHPL_0 ($b_{nom} = 2\,\mathrm{m}$)

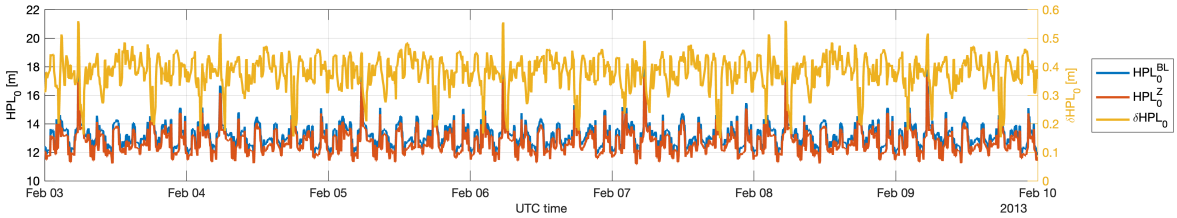
Figure 6.2: Global contours of the reduction of computed fault-free HPL using the zonotope-based method compared to the baseline ARAIM algorithm (maximum and minimum values). Two scenarios concerning the predefined nominal bias are simulated: $b_{nom} = 0.75\,\mathrm{m}$ and $b_{nom} = 2\,\mathrm{m}$.



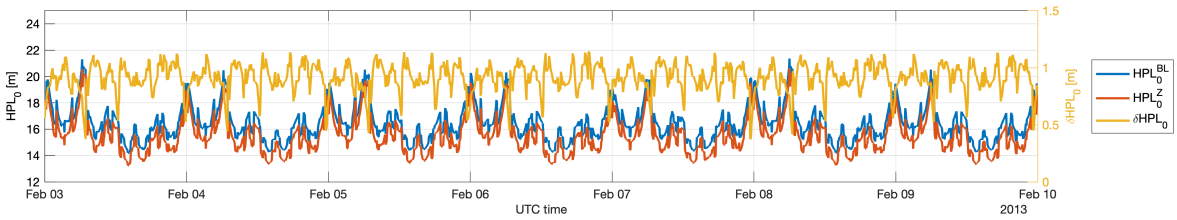
(a) Latitude 0°, Longitude 10° ($b_{nom} = 0.75\,\mathrm{m}$)



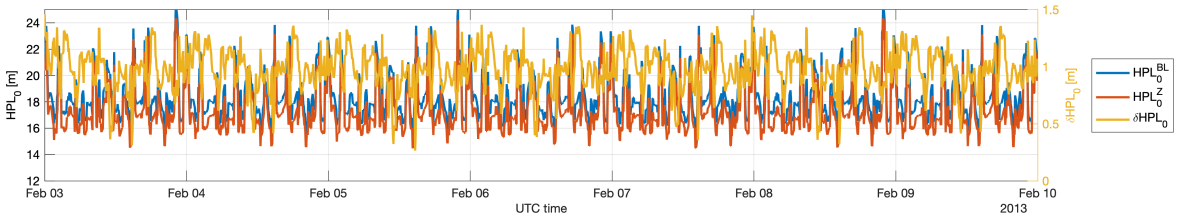
(b) Latitude 50°, Longitude 10° ($b_{nom} = 0.75\,\mathrm{m}$)



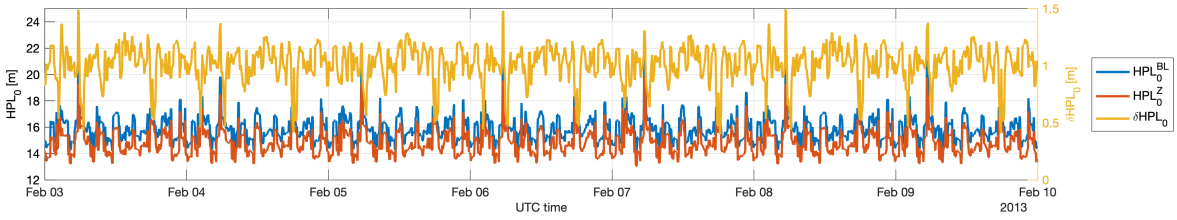
(c) Latitude 80°, Longitude 10° ($b_{nom} = 0.75\,\mathrm{m}$)



(d) Latitude 0°, Longitude 10° ($b_{nom} = 2\,\mathrm{m}$)



(e) Latitude 50°, Longitude 10° ($b_{nom} = 2\,\mathrm{m}$)



(f) Latitude 80°, Longitude 10° ($b_{nom} = 2\,\mathrm{m}$)

Figure 6.3: Time series of fault-free HPL simulations: Baseline ARAIM (HPL_0^{BL}, blue), zonotope-based method (HPL_0^Z, red), and achieved reduction (δHPL_0, yellow). Two nominal biases are considered: $b_{nom} = 0.75, \mathrm{m}$ (a–c) and $2, \mathrm{m}$ (d–f), across three locations: Latitude 0°, Longitude 10° (low-latitude), Latitude 50°, Longitude 10° (mid-latitude) and Latitude 80°, Longitude 10° (high-latitude).

In conclusion, the refined HPL computation scheme proves advantageous in comparison with the baseline ARAIM algorithm, generating smaller fault-free HPL for aviation users. This improvement can be achieved at a global scale, depending on satellite availability and the predetermined nominal biases.

6.2 Interval-based integrity monitoring for urban navigation

This section aims to evaluate the interval-based integrity monitoring approach developed in Chap. 5, which is equipped with the distribution-free uncertainty modeling outlined in Chap. 3, within the context of urban navigation. The kinematic and static datasets used are introduced in Sec. 6.2.1, followed by the experimental setup and Integrity Support Parameters (ISP) configurations in Sec. 6.2.2. Sec. 6.2.3 introduces threat modeling, addressing nominal error bounding through interval and Gaussian overbounding methods, as well as handling unknown local threats. The performance of fault-free error bounding is analyzed in Sec. 6.2.4, followed by an analysis of the loss of integrity (LOI) in Sec. 6.2.5. Finally, Sec. 6.2.6 examines the impact of protection functions against unknown local threats.

6.2.1 Data collection

Two datasets are utilized: one in the kinematic mode collected in urban environments and the other in the static mode collected at a continuously operating GNSS reference station.

Kinematic dataset

The kinematic measurement dataset was collected by a moving vehicle in Hannover, Germany, on Dec. 9th, 2021, over approximately 2.5 hours, as shown in Fig. 6.4 and detailed information in Tab. 6.2. The vehicle was equipped with a *Septentrio PolaRx5e* receiver and a *navigation*-grade Inertial Measurement Unit (IMU) *iMAR iPRENA-I-A*. Post-processed GNSS/IMU tightly-coupled solutions were generated using the *TerraPos* software with a centimeter-level precision, serving as the reference trajectory. The dataset is publicly available at https://doi.org/10.25835/75o9yrc0 (Axmann et al., 2023).

The dataset was initially sampled at 10 Hz. For computational efficiency, it is resampled to 1 Hz, which is appropriate given the low driving speeds in urban areas.

Table 6.2: Main information for the kinematic and static datasets.

Parameter	Kinematic dataset	Static dataset
Sampling interval	1 sec	30 sec
Total number of epochs	9866	20160
Trajectory length	Around 17 km	N/A
Approximated coordinates (latitude, longitude)	Starting point: $(52.3865°N,\ 9.7156°E)$	$(52.3879°N,\ 9.7125°E)$

Static dataset

The static measurement dataset was collected at a continuously operating GNSS reference station, labeled as *EE01*, which is situated on the rooftop of the HiTec building on the campus

Figure 6.4: Data collection: the permanent station collecting the static dataset and the trajectory of the kinematic dataset are marked in the map as the light blue dot and dark blue lines, respectively.

of Leibniz University Hannover, Germany (see Fig. 6.4). A *Septentrio PolaRx5TR* receiver, which belongs to the same series as the receiver used for the kinematic dataset, was used to maintain similar performance levels. This receiver operated continuously for one week, overlapping with the period during which the kinematic dataset was collected.

6.2.2 Experimental setup

Integrity support parameters

Currently, there are no widely accepted standards for urban navigation integrity. To conduct a meaningful evaluation, the IR requirement is set to 10^{-4}. This setting is not tied to any specific application but is based on the volume of data collected: Analyzing navigation integrity over nearly 10^4 epochs (kinematic dataset) and 2×10^4 epochs (static dataset) necessitates an IR requirement of 10^{-4} as a baseline to demonstrate the evaluation's effectiveness.

It is noteworthy that this setting aligns with the recommended requirements for payment-critical and regulatory-critical road applications, as outlined in the Report on Road & Automotive User Needs and Requirements (EUSPA, 2021).

The remaining ISPs are listed in Tab. 6.3. For the real-time integrity monitoring of the kinematic dataset, the satellite fault probability (P_{sat}) is set to 10^{-4}, which is further discussed in Sec. 6.2.3. Moreover, this probability is set to zero for static fault-free analysis. The constellation fault probability (P_{const}) is not considered in this work, and the threshold for not-monitored probability (P_{thres}) is set slightly lower than the IR requirement, consistent with typical ARAIM settings in the literature such as Blanch et al. (2015a).

Given the ISPs, it is sufficient to monitor only single-measurement faults. This means that the maximum number of monitored faults is one ($r = 1$), so fault modes with more

Table 6.3: Integrity support parameters (ISP) in the urban navigation experiment.

Parameter	Value
IR requirement (I_{REQ}), in which	10^{-4}
for the horizontal component ($I_{REQ,H}$)	0.5×10^{-4}
Satellite-wise fault probability (P_{sat})	0 (for static fault-free analysis) 10^{-4} (for kinematic real-time monitoring)
Constellation-wise fault probability (P_{const})	not considered
Threshold for not-monitored probability (P_{thres})[1]	4×10^{-5}
CR requirement (C_{REQ})[2]	10^{-4}
Threshold for iterative computations of PL (TOL_{PL})[3]	10^{-2}

[1] P_{thres} is necessary for the multiple-hypothesis framework, applied to the interval approach and ARAIM.
[2] C_{REQ} is used in ARAIM and RB RAIM.
[3] TOL_{PL} is only used in ARAIM for computing the PL.

than one simultaneous fault are accounted for by the "not-monitored" probability P_{NM} (see Eq. 5.55). For example, based on the formulas in Sec. 5.2.2, P_{NM} for $r = 1$ varies from 2.10×10^{-7} to 7.78×10^{-6} with 7 to 40 measurements. This range is lower than the threshold ($P_{thres} = 4 \times 10^{-5}$). This setup enables the legacy RB RAIM to be performed for comparison.

Processing methods and corrections applied

Two modes of pseudorange-based Single Point Positioning (SPP) are analyzed: (i) single-frequency SPP with broadcast correction products and (ii) dual-frequency SPP with precise IGS orbit and clock products. Tab. 6.4 summarizes the applied error correction methods.

Table 6.4: Processing methods and corrections applied.

	Positioning mode	
	Single-frequency SPP	Dual-frequency SPP
Observation types (RINEX naming format)	GC1C, RC1C, EC1C	GC1C/GC2W, RC1C/RC2C, EC1C/EC5Q
Sat. orbit & clock error	Broadcast ephemerids	IGS WUM rapid products
Ionospheric delay	GPS Klobuchar model	Ionosphere-free LC
Tropospheric delay	Saastamoinen model with ISO2533 standard atmosphere	
Satellite group delay	Broadcast SGD	WUM OSB products
Multipath error/NLOS	Not corrected	

6.2.3 Threat modeling

Introduction

The threat model in the experiment is defined on the basis of ARAIM threat model (Sec. 2.4.2) with special adaptations to reflect the unique challenges of urban environments. The following aspects are considered:

- For the **receiver hardware noise**, the uncertainty specific to the receiver type used in the experiment is analyzed.
- **Multipath effect and NLOS** are dominant local threats in urban environments. While single-reflected multipath effect can be handled by MEE-based interval bounding discussed in Sec. 3.4, more complex multipath effect and NLOS conditions are difficult to predict, or model precisely. Therefore, these are considered as narrow faults in this work.

Nominal error bounding

The following nominal error sources are considered:

- Signal-In-Space Range Error (SISRE) due to orbit and clock errors
- Residual ionospheric errors
- Residual tropospheric errors
- Multipath errors (single-reflected or refracted)
- Receiver hardware noise
- Nominal signal deformation

Other nominal error sources are either corrected to a significant level or considered negligible.

The strategies used for bounding these nominal errors in the two positioning modes are summarized in Tab. 6.5. Specifically, the subsequent subsection details how receiver noise is treated in this experiment.

Table 6.5: Nominal error bounding approaches in the experiment.

Nominal errors	**Positioning mode**	
	Single-frequency SPP	Dual-frequency SPP
Sat. orbit & clock error	Gaussian overbounds based on URA specification	Gaussian distribution based on IGS product quality assessment[1]
Res. tropospheric error	Interval bounds (Sec. 3.2)	
Res. ionospheric error	Interval bounds (Sec. 3.3)	Not considered
Multipath error	Interval bounds (Sec. 3.4)	
Receiver code noise	Empirically determined Gaussian overbounds	
Nom. signal deformation	Not considered	Interval bounds[2]based on worst-case nominal bias b_{nom}

[1] We choose the reported RMS value (rounded to 3 cm) for IGS final precise orbit products as the sigma of the assumed Gaussian distribution (Johnston et al., 2017).
[2] We take the nominal bias value from ARAIM $b_{nom} = 0.75$ m to define symmetric interval radii for the uncertainty due to nominal signal deformation.

The RAIM approaches are implemented for comparative analysis. The residual tropospheric and ionospheric errors are handled stochastically (refer to Tab. 5.4). Meanwhile, the multipath error is treated using the MOPS airframe multipath model, which is included in the baseline ARAIM algorithms (Blanch et al., 2015a). The model is represented by the following equation:

$$\sigma_{MP}^2 = 0.13 + 0.53 \times e^{\frac{\theta}{10}} \, [\mathrm{m}^2], \tag{6.5}$$

where θ is the satellite elevation angle in degrees. This model, designed for GPS signals, is also applied to GLONASS and Galileo signals in the experiments due to the lack of corresponding standard specifications.

Receiver hardware noise

The receiver hardware noise is modeled in a stochastic manner. Precisely, an overbounding Gaussian distribution for the receiver hardware noise is estimated to define the measurement intervals. While various stochastic models exist in the literature, such as Hartinger and Brunner (1999) and Luo et al. (2009), the following elevation-dependent model is chosen:

$$\sigma(\theta)^2 = \left(\frac{\sigma_0}{\sin(\theta)}\right)^2, \tag{6.6}$$

where θ is the satellite elevation and $\sigma(\theta)$ represents the corresponding standard deviation, describing the uncertainty due to the receiver hardware noise. σ_0 is a pre-determined overbounding sigma value, which is input into the integrity monitoring algorithm.

It is important to note that other models could also be used; however, the tightness of bounding (i.e., the margin between the bounds and actual errors) would vary, eventually affecting the estimation error bounding performance, as discussed in Sec. 4.4.3. In this work, we select the model in Eq. 6.6 based on expert knowledge and empirical experience, e.g., the normalized experimental distributions for the elevations of interest tend to exhibit Gaussian behavior for most GNSS signals.

To estimate σ_0, the one-week static data from *EE01* station is analyzed. This is an open-sky measurement environment and, hence, is dedicated to simulating a fault-free scenario. Prior to analysis, several aspects were checked to confirm the fault-free assumption:

- **Constellation Service Provider (CSP)s' announcements** including *Notice Advisory to Navstar Users*, *Notice Advisory to GLONASS Users*, and *Notice Advisory to Galileo Users* were reviewed for potential satellite usability. The following events were identified:
  - GPS PRN02 unusable from 13:04 until 19:34 on Dec. 9th, 2021 (GPS time);
  - GPS PRN22 unusable for the entire week (GPS time);

  The CSPs of GLONASS and Galileo did not report any unusability alerts during the experiment.
- **3D city model-based ray tracing** to identify potential blocked or severely affected signals. Fig. 6.5 showcases the sky visibility for the *EE01* station, indicating that a small portion of direct signals may be blocked. Any signals from this area of the blocked sky are excluded from fault-free analyses.

After excluding all potentially *faulted* signals, the Multipath Linear Combination was formed (cf. Eq. 3.27). The multipath error term is assumed to be zero, and hence, the receiver code noise is considered to dominate the LC values, resulting in random variations, as illustrated in Fig. 6.6 and Fig. 6.7 (the second column with respect to each code observable).

The overbounding sigma σ_0 is estimated from normalized LC residuals based on Eq. 6.6. The first columns of Fig. 6.6 and Fig. 6.7 present the folded CDFs of the normalized residuals and corresponding overbounding Gaussian distributions for GPS, GLONASS and Galileo code observables. Resulting σ_0 values are summarized in Tab. 6.6.

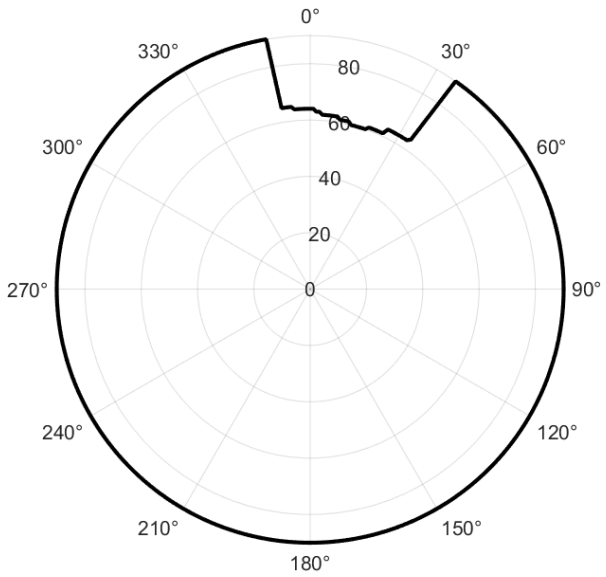


Figure 6.5: Sky visibility of the *EE01* station: one building in front of the station can block a small portion of the direct signals.

The impact of potential multipath-affected or NLOS signals is considered negligible for the following reasons:

- Significant outliers observed in the Multipath LC are excluded from the analysis to mitigate the influence of such signals;
- The determined overbounding sigma accounts for the effect of remaining multipath errors as part of the nominal error budget.

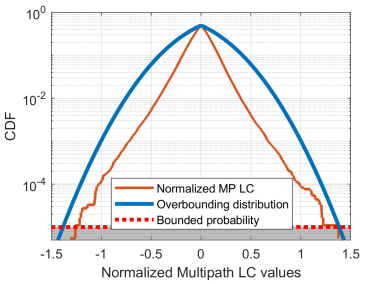
Table 6.6: Overbounding standard deviations derived from the normalized Multipath LC values.

LC	**Code observable**	**Phase observables**	**Overbounding sigma value** [m]
1	GC1C	GL1C, GL2W	0.0.3244
2	GC2W	GL2W, GL1C	0.1334
3	RC1C	RL1C, RL2C	0.6163
4	RC2C	RL2C, RL1C	0.4521
5	EC1C	EL1C, EL5Q	0.2709
6	EC5Q	EL1C, EL5Q	0.1320

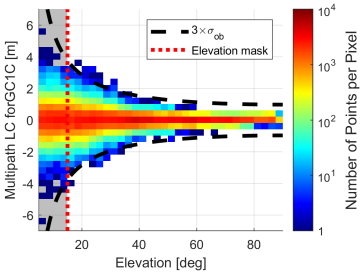
Unknown local threats

Local threats, such as NLOS, pose significant challenges for GNSS urban navigation. These threats are difficult to correct or compensate for using GNSS augmentation information (e.g., SBAS). Additionally, they are not easily captured by nominal error models or statistical frequencies, which are necessary to navigation integrity monitoring (Crespillo et al., 2022). Consequently, the navigation system may experience degraded integrity performance or, in extreme cases, loss of integrity and continuity.

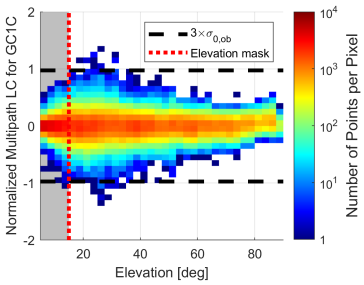
Functions targeted at protecting the system against the presence of local threats are essential. Possible solutions include various interference detection methods and hardware-based interference mitigation techniques. This experiment considers three typical approaches:



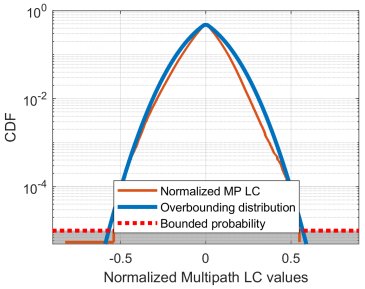
(a) Folded CDF for normalized MP LC (GC1C, GL1C, GL2W)



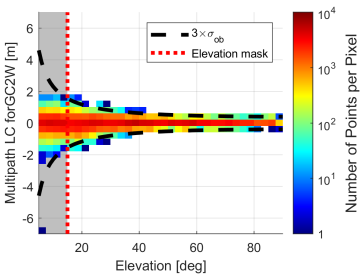
(b) MP LC w.r.t elevation (GC1C, GL1C, GL2W)



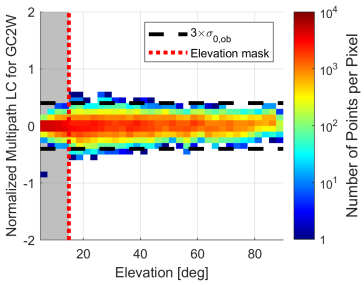
(c) Normalized MP LC w.r.t elevation (GC1C, GL1C, GL2W)



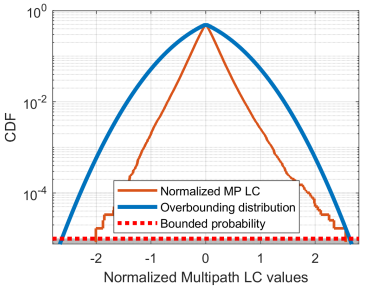
(d) Folded CDF for normalized MP LC (GC2W, GL2W, GL1C)



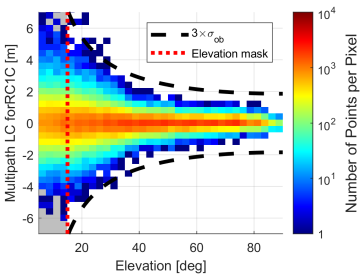
(e) MP LC w.r.t elevation (GC2W, GL2W, GL1C)



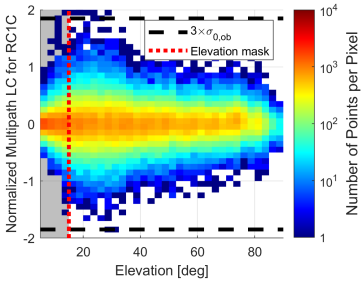
(f) Normalized MP LC w.r.t elevation (GC2W, GL2W, GL1C)



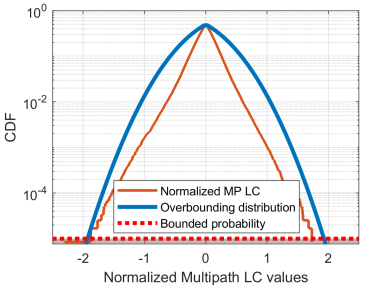
(g) Folded CDF for normalized MP LC (RC1C, RL1C, RL2C)



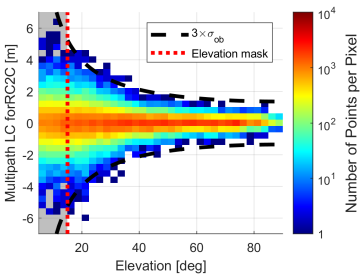
(h) MP LC w.r.t elevation (RC1C, RL1C, RL2C)



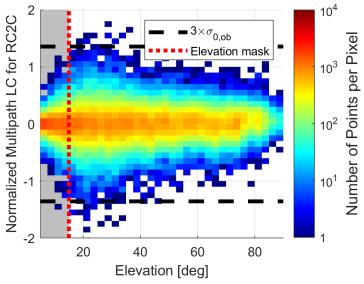
(i) Normalized MP LC w.r.t elevation (RC1C, RL1C, RL2C)



(j) Folded CDF for normalized MP LC (RC2C, RL2C, RL1C)



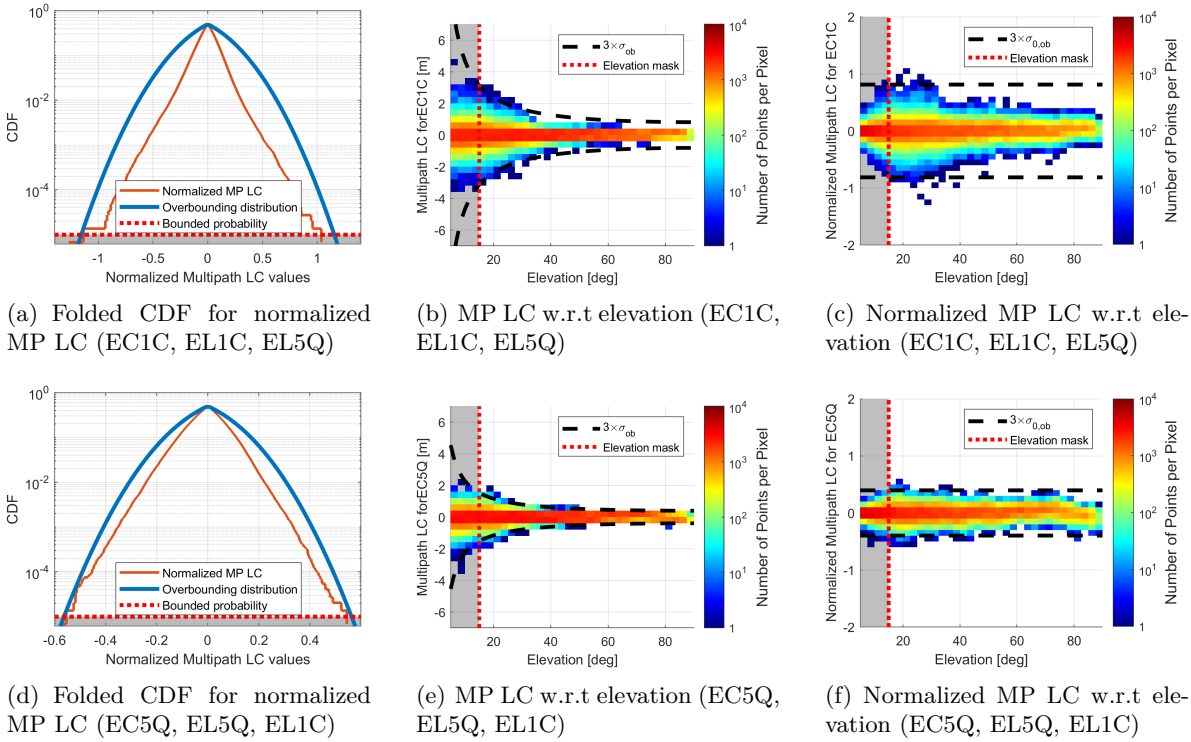
(k) MP LC w.r.t elevation (RC2C, RL2C, RL1C)



(l) Normalized MP LC w.r.t elevation (RC2C, RL2C, RL1C)

Figure 6.6: Forming Multipath LC to derive overbounding standard deviations for receiver hardware noise with respect to GPS and GLONASS code observables. Observations with elevation lower than 15° are not involved in the statistics (highlighted in gray areas); the indented overbounding Gaussian distributions bound the observations' CDF down to 10^{-5}.

- **Receiver-level interference mitigation.** The *Septentrio PolaRx5* receiver enables the function of *A-Posteriori Multipath Estimation* (APME), using extra correlators in each tracking channel to estimate and correct the multipath error in pseudorange and carrier phase measurements.



(a) Folded CDF for normalized MP LC (EC1C, EL1C, EL5Q)

(b) MP LC w.r.t elevation (EC1C, EL1C, EL5Q)

(c) Normalized MP LC w.r.t elevation (EC1C, EL1C, EL5Q)

(d) Folded CDF for normalized MP LC (EC5Q, EL5Q, EL1C)

(e) MP LC w.r.t elevation (EC5Q, EL5Q, EL1C)

(f) Normalized MP LC w.r.t elevation (EC5Q, EL5Q, EL1C)

Figure 6.7: Forming multipath LC to derive overbounding standard deviations for receiver hardware noise with respect to Galileo code observables. Observations with elevation lower than 15° are not involved in statistics (highlighted in gray areas); the indented overbounding Gaussian distributions bound the observations' CDF down to 10^{-5}.

- **Signal classification using 3D city model-based ray tracing.** The experiment employs MEE-based interval bounding to address multipath errors in signals that are classified as multipath-affected.
- **Data editing.** Excluding measurements based on reasonable checks and criteria can reduce significant measurement errors. This process is referred to as data editing by Crespillo et al. (2022). This work applied (i) masks for C/N0 value and satellite elevation:
  - C/N0 mask: 25 dB-Hz (applied to GS1C, RS1C and ES1C observables);
  - Elevation mask: 15°.

  and (ii) NLOS/multipath detection based on Multipath LC: A detection criterion is set to $k_{MP} \cdot \sigma(\theta)$, where the coefficient $k_{MP} = 3.89$ is determined based on the tail probability 10^{-4} to be protected, as discussed in the subsequent subsection, and $\sigma(\theta)$ is calculated using Eq. 6.6). Multipath LC observations are obtained using Eq. 3.27 with procedures introduced in Sec. 3.4.3. Observations greater than the threshold are excluded from estimations.

Excessive protection against local threats may significantly impact satellite availability, potentially leading to insufficient measurements for position estimation and integrity monitoring functions. Therefore, a trade-off must be sought without harming system availability. Sec. 6.2.3 will explore its impact.

Fault probabilities

The fault probabilities associated with local threats are yet to be fully understood. In the experiment, a satellite-wise fault probability $P_{sat} = 10^{-4}$ is approximated for every satellite based on the following considerations:

- P_{sat} in ARAIM threat models does not exceed this level for GPS, GLONASS and Galileo satellites (see Tab. 2.4);
- Most local threats are expected to be excluded after applying the protection functions;
- The overbounding Gaussian distributions for nominal receiver hardware noise are valid down to 10^{-5}, supporting this setting.

However, the second assumption may be questionable. On the one hand, it is difficult to identify local threats with GNSS-only measurements, especially in challenging environments where a majority of measurements are affected; on the other hand, a balance must be considered between ensuring satellite visibility and avoiding disruptions to navigation functions.

6.2.4 Fault-free error bounding analysis

For applications such as infrastructure-based augmentation systems, e.g., SBAS, the PL computation must account for the nominal error budget while the augmentation system monitors fault events (anomalies). In this context, exploring the fault-free HPL is important.

This section utilizes the static dataset for fault-free analysis, excluding all potential faulted signals as discussed in Sec. 6.2.3. Under the fault-free assumption:

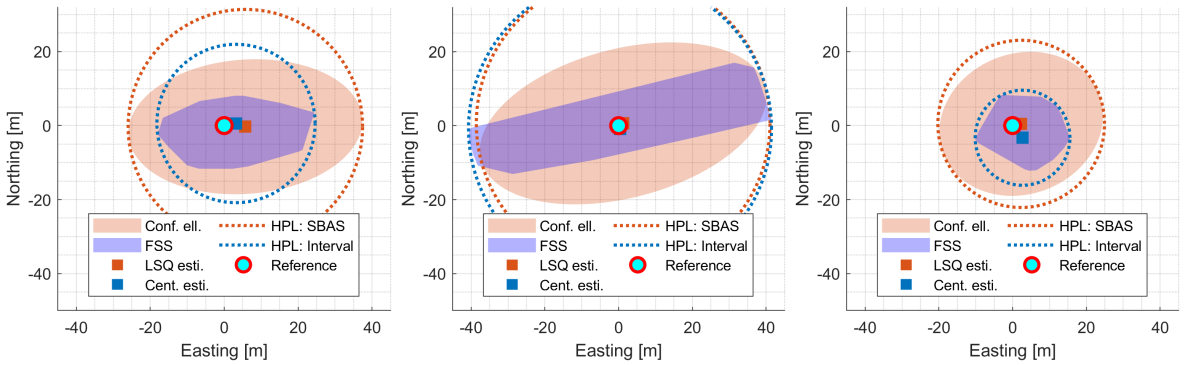
- The FSS for the interval-based approach is in the form of polytopes (cf. Sec 4.3.1), with the point solution and HPL determined from the Chebyshev center and radius.
- The least-squares solution is expected to be protected by a MOPS SBAS-style HPL with the detailed formula provided in Annex. B.

Fig. 6.8 presents three example epochs for the interval and stochastic approaches:

- **Example 1** and **Example 3** show cases where the HPL from the interval approach is smaller than the SBAS-style HPL. The central estimator's performance varies, being more accurate in Example 1 but less in Example 3 compared to the least-squares solutions.
- **Example 2** illustrates a scenario with poor satellite geometry, where the interval-based HPL is slightly larger than the SBAS-style HPL (approximately 41.1 m and 39.8 m, respectively), where both point solutions remain close to the true position (low HPE, the central estimated solution is slightly more accurate, approximately 0.9 m and 1.4 m, respectively).

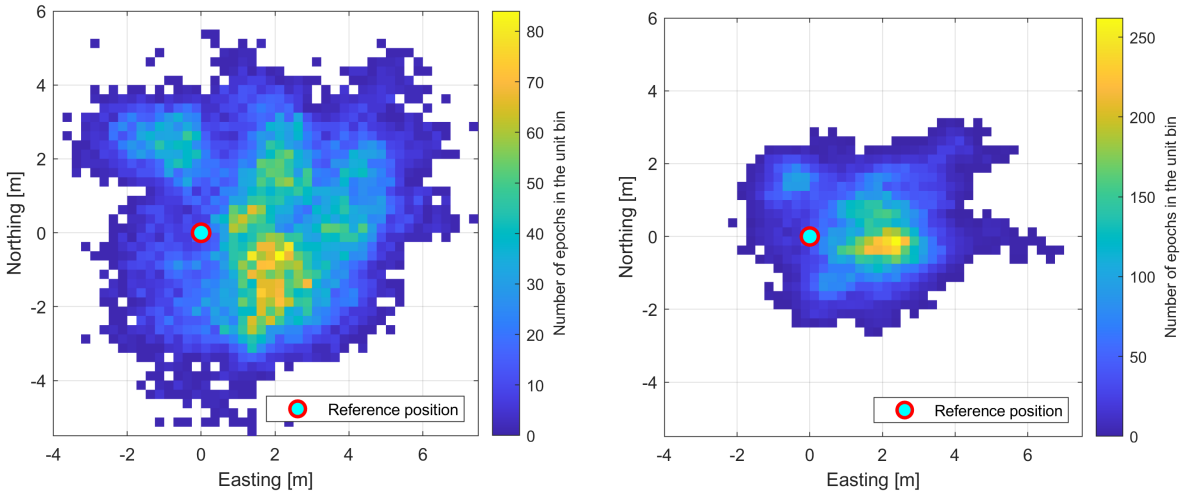
Fig. 6.9 compares the scatter density of all estimated solutions, i.e., the central estimator (Fig. 6.9(a)) and the least-squares estimator (Fig. 6.9(b)). The central estimated solutions exhibit a wider spread than the least-squares solutions.This is further reflected in the time series of HPE (Fig. 6.10(a)), where the central estimator generally shows greater HPE than the LSQ estimator. The ECDF in Fig. 6.10(c) also confirms this observation.

In contrast, the interval-based approach provides significantly lower HPL values than the SBAS-style HPL, as shown in the time series in Fig. 6.10(a) and ECDF plot in Fig. 6.10(d). The reduction in HPL is generally up to 10 m, with a trade-off of approximately 1 m in accuracy.



(a) Example 1: Dec. 6th, 10:59:30 (b) Example 2: Dec. 9th, 19:41:30 (c) Example 3: Dec. 11th, 03:35:30

Figure 6.8: Three fault-free example epochs for estimators and error bounds: the interval approach (central estimator & feasible solution set) and the stochastic approach (least-squares estimator & confidence ellipse).



(a) Central estimator (b) Least-squares estimator

Figure 6.9: Scatter density of fault-free positioning for the central estimator and the least-squares estimator.

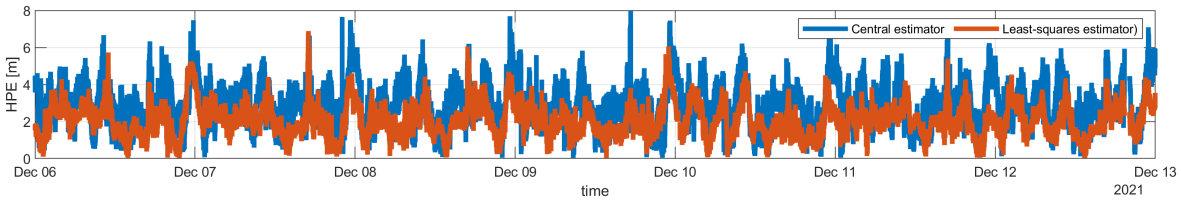
This trade-off does not precisely align with the simulations discussed in Sec. 4.4.3, due to the following reasons:

- The actual error components are more complex than simulated scenarios;
- The developed observation error bounds are more conservative than simulated scenarios, i.e, loose bounds, which can highly affect the position-domain performance.

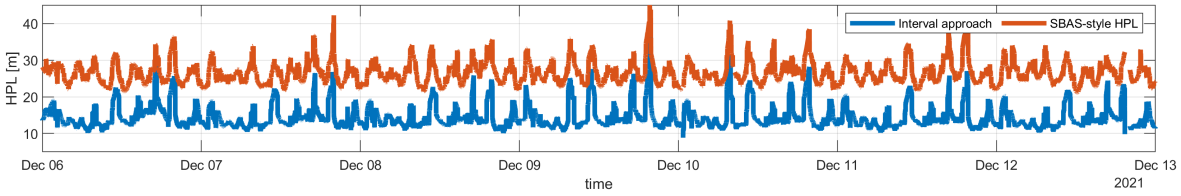
Conclusively, the interval-based approach yields significantly tighter position bounds with integrity assurance under fault-free conditions, offering higher system availability compared to the stochastic approach. This makes it particularly advantageous for applications such as infrastructure-based augmentation systems.

6.2.5 Comparative analysis for loss of integrity

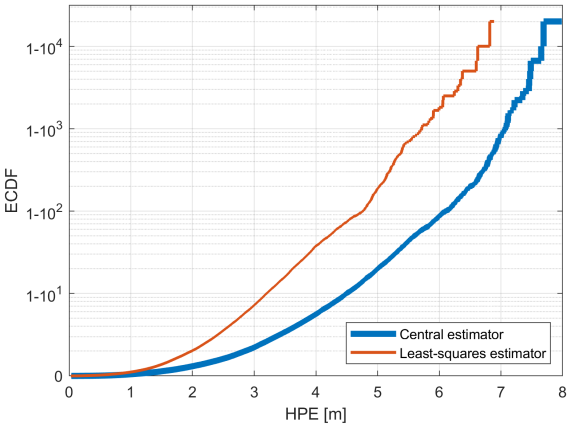
This section evaluates the integrity performance of the interval-based approach, with a special focus on fault detection (FD), and compares it with the classical RB RAIM and the state-of-the-art baseline SS ARAIM algorithm. Both single-frequency and dual-frequency SPP modes



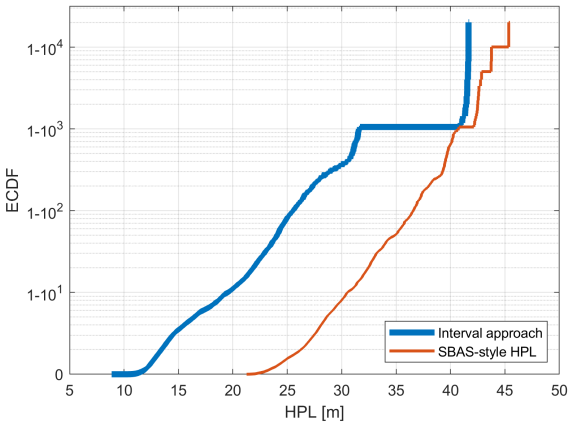
(a) Fault-free HPE times series



(b) Fault-free HPL times series



(c) Fault-free HPE ECDF



(d) Fault-free HPL ECDF

Figure 6.10: Fault-free HPE/HPL using different methods for the static dataset in the form of (a-b) time series, and (c-d) ECDF.

are analyzed, as configured in Tab. 6.4, with corresponding nominal error modeling shown in Tab. 6.5.

Single-frequency SPP

For single-frequency SPP, we compare the interval-based approach, RB RAIM and SS ARAIM using GPS-only (G) and GPS+GLONASS+Galileo (GRE) constellations. Two key metrics are considered:

- **MI occurrence.** Ideally, the MI occurrence should remain below the IR requirement;
- **HPL values.** Lower HPL indicates better availability performance. For this LOI analysis, HPL is calculated only when an alert is not triggered (no fault detection). If an alert is issued, the user is notified of system unavailability, and HPL is neither calculated nor provided to the user, following the procedures outlined in Sec. 5.2.3.

Fig. 6.11 shows the ECDF of HPL across all scenarios, and the results are further visualized using "partial" Stanford diagrams (i.e., without specified AL) in Fig. 6.12.

Notably, no MI occurrence was observed, indicating full protection in terms of loss of integrity. In terms of availability:

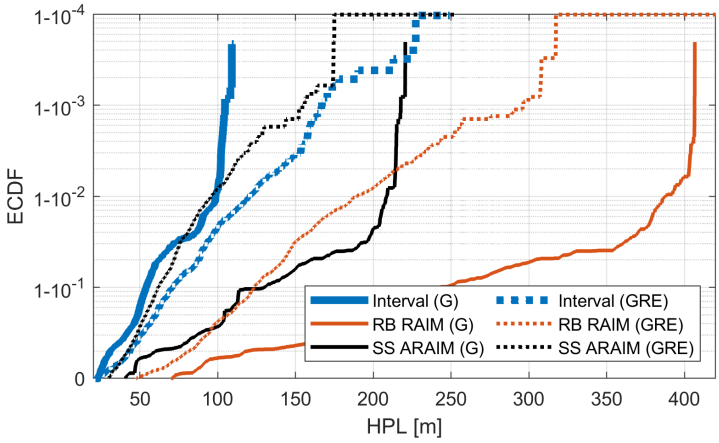


Figure 6.11: ECDF of HPL for single-frequency SPP.

- **GPS-only (G)**: The interval-based approach provides lower HPL values compared to the other two methods, consistent with the simulation analysis for IR bounds in Sec. 5.3.
- **Three constellations (GRE)**: SS ARAIM generally achieves the tightest HPL values. The interval approach, while slightly looser, still outperforms RB RAIM. The 68.27%, 95.45%, 99.73% percentiles of the HPL values are summarized in Tab. 6.7.

Table 6.7: Integrity and availability performance for the three approaches in single-frequency SPP.

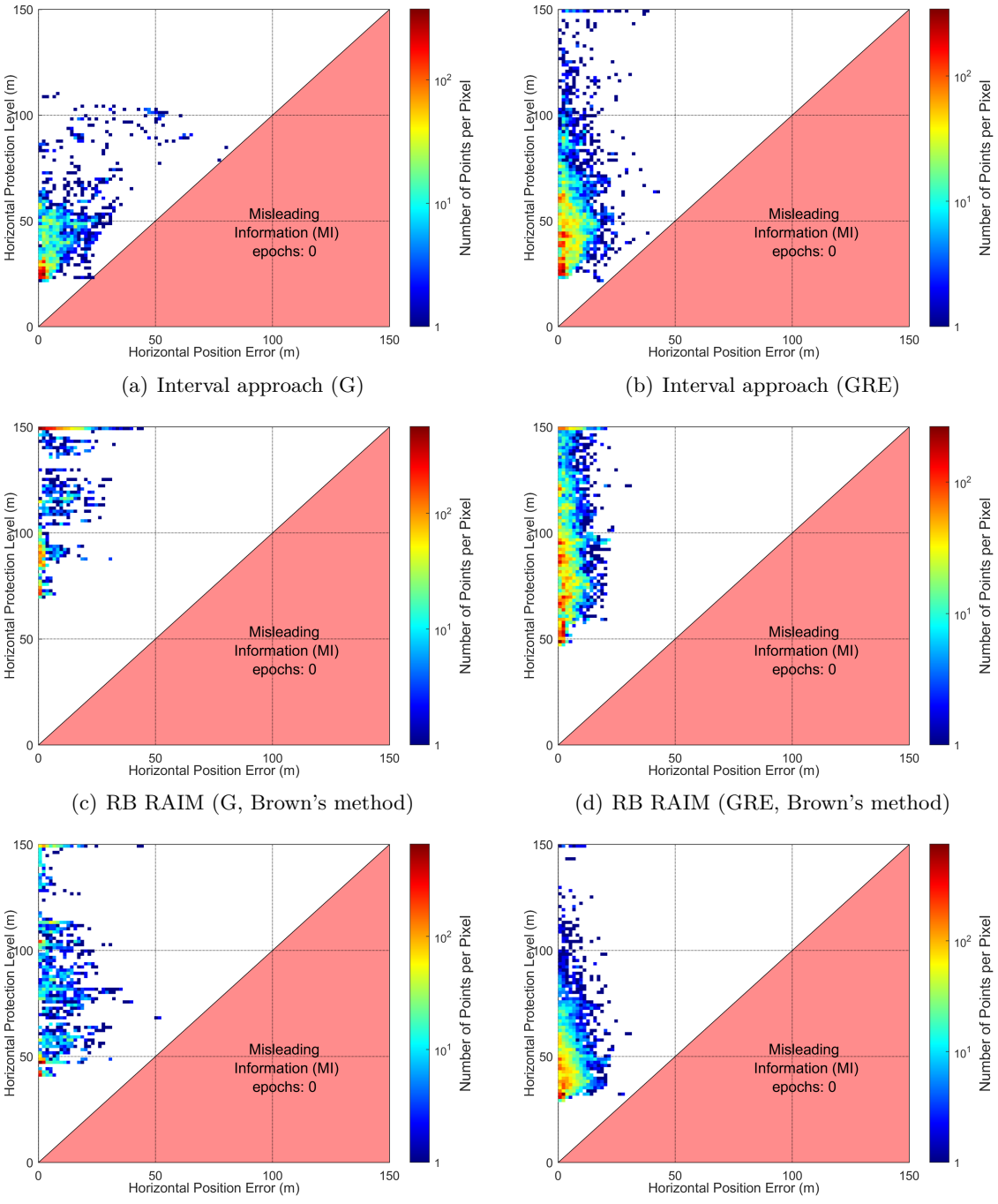
Parameters	Interval approach		RB RAIM		SS ARAIM	
	G	GRE	G	GRE	G	GRE
MI occurrence	0	0	0	0	0	0
HPL (68.27%, [m])	42.4	51.8	177.6	94.6	88.9	48.5
HPL (95.45%, [m])	67.2	88.6	320.9	144.0	169.7	72.5
HPL (99.73%, [m])	114.8	154.2	404.4	240.8	214.8	121.1

This trend differs from the simulation analysis in Sec. 5.3.3, reflecting the real-world performances of the three approaches. The discrepancy mainly stems from their differences in PL computation and risk allocation. For example, RB RAIM applies a more conservative calculation of position bounds (HPL, see Sec. 2.4.5), leading to the worst availability performance. The interval approach, with moderate conservatism, offers superior availability in single-constellation scenarios and remains competitive when multiple constellations are used.

These findings underscore a critical insight for designing integrity algorithms: The tightness of analytical IR bounds is fundamental for system availability. However, in real-world applications, availability is also significantly affected by how the risk domain is translated into the position domain. This includes PL calculation and the allocation of risk across multiple states, both of which should balance the computational efficiency and effectiveness to enhance availability.

Dual-frequency SPP

This section examines dual-frequency SPP, which gains more accuracy due to the formed ionosphere-free LC and the use of IGS precise orbit and clock products. The evaluation focuses on three metrics:



(a) Interval approach (G)

(b) Interval approach (GRE)

(c) RB RAIM (G, Brown's method)

(d) RB RAIM (GRE, Brown's method)

(e) SS ARAIM (G, Baseline algorithm)

(f) SS ARAIM (GRE, Baseline algorithm)

Figure 6.12: Partial Stanford diagrams for single-frequency SPP using GPS-only (G) and GPS+Galileo+GLONASS (GRE) signals.

- **HMI occurrence**, the primary performance metric, evaluated against the specified AL. This analysis sets the AL to 25 m to meet the requirements for payment-critical and regulatory-critical applications as outlined by EUSPA (2021). **HMI occurrence** should in no case exceed the IR requirement.
- **MI occurrence**, consistent with the evaluation for single-freqeuncy SPP.
- **System availability**, defined as the percentage of epochs where no alert is declared (i.e., no FD) and the computed HPL does not exceed AL.

In certain scenarios, post-FD performance may be relevant, particularly when position bounds need to accommodate the impact of detected faults while maintaining continuous operation. For instance, position bounds might still meet operational requirements despite detected but unresolved faults. The interval-based approach addresses this by inflating measurement intervals and recalculating the FSS (cf. Sec. 5.2.3). For RAIM approaches, a Horizontal Uncertainty Level (HUL) is computed using actual measurements. This study adopts Young's method for RB RAIM (Young et al., 1996) and the method in WG-C ARAIM TSG (2019) for SS ARAIM. The detailed formulas are provided in Annex. B. The evaluation includes both alerted and non-alerted epochs, i.e., the system is considered *available* as long as the computed HPL (for interval approach) or HUL (for RAIM approaches) does not surpass the specified AL.

Tab. 6.8 summarizes the results, and Fig. 6.13 illustrates them in Stanford diagrams.

Table 6.8: Integrity and availability performance for the three approaches in dual-frequency SPP.

Parameters	**Interval approach**		**RB RAIM**		**SS ARAIM**	
	no alert	all	HPL (no alert)	HUL	HPL (no alert)	HUL
HMI occurrence	0	0	0	0	0	8
MI occurrence	28	59	0	9	11	48
Availability [%]	73.8	85.7	48.1	84.9	61.3	92.4

Key observations are summarized as follows:

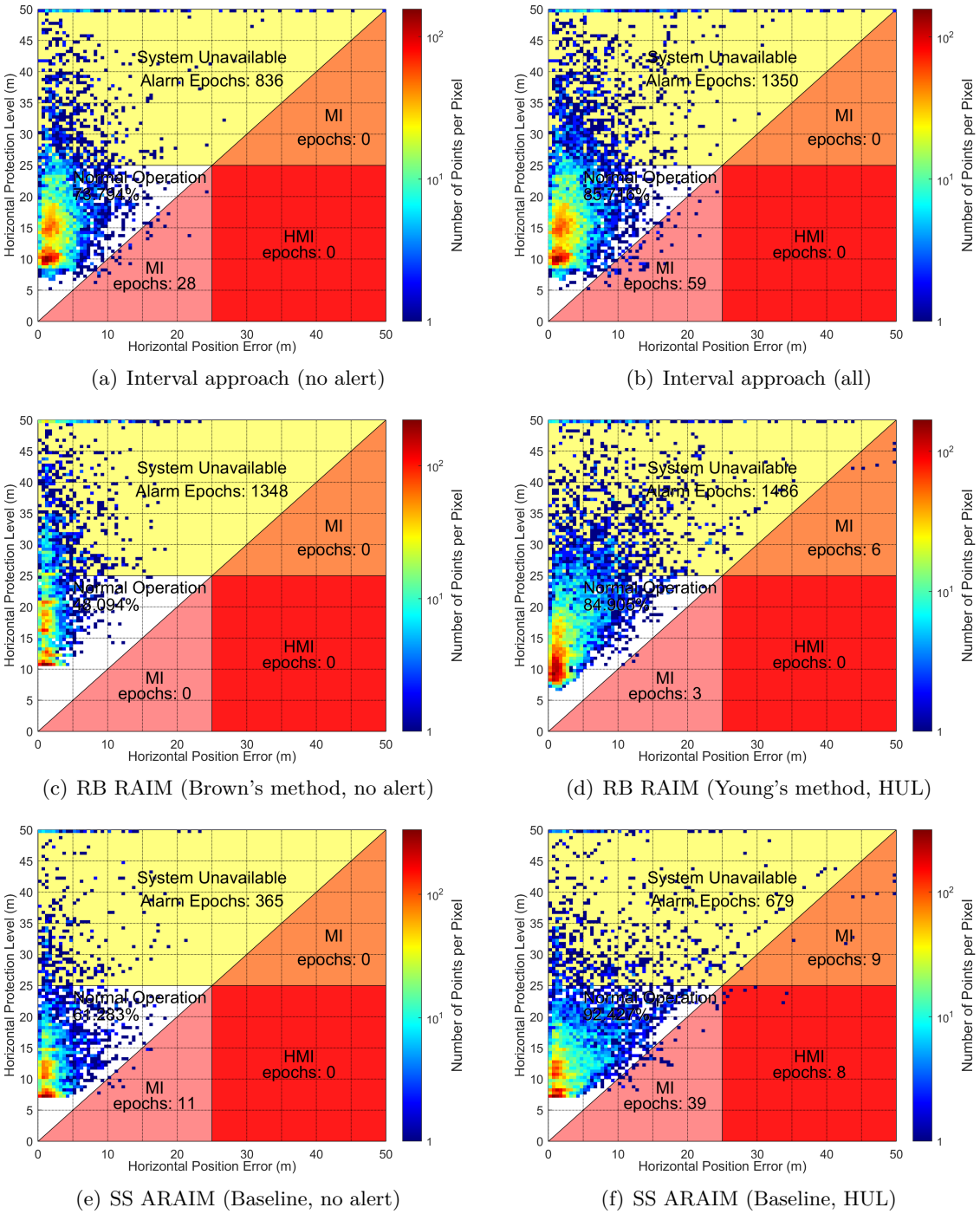
- **Non-alerted epochs:** All approaches provide full protection against HMI occurrence concerning the specified AL. Notably, RB RAIM, while demonstrating zero MI occurrences, has the lowest availability due to the highest HPL values. The interval approach exhibits significantly better availability (73.8%) compared to the SS ARAIM (61.3%) and RB RAIM (48.1%).
- **All epochs (alerted and non-alerted):** The interval approach and RB RAIM remain to guarantee protections against HMI, both reporting zero HMI occurrences. In contrast, the SS ARAIM experiences eight HMI occurrences, significantly exceeding the IR requirement ($8 \times 10^{-4} > 0.5 \times 10^{-4}$, for horizontal). Violating the primary safety concerns renders any availability advantage meaningless. Despite a higher number of MI occurrences, the interval approach maintains slightly better availability than RB RAIM (85.7% and 84.9%, respectively).

Overall, the developed interval-based approach demonstrates robust integrity performance compared to the RAIM approaches in the dual-frequency SPP, with a notable advantage in availability, particularly in non-alerted scenarios. However, further refinements are needed to address MI occurrences. A potential solution is enhancing the protection functions against unknown local threats, discussed in the subsequent section.

6.2.6 Impact of protection functions against local threats

Sec. 6.2.3 introduces various possible protection functions against unknown local threats. The choice of using a single function or a combination of multiple functions can influence both integrity and availability. This section aims to evaluate the impact of the following approaches:

- **C/N0 mask** (25 dB-Hz). Signals with lower C/N0 values are excluded from estimation (applied to GS1C, RS1C and ES1C observables);



(a) Interval approach (no alert)

(b) Interval approach (all)

(c) RB RAIM (Brown's method, no alert)

(d) RB RAIM (Young's method, HUL)

(e) SS ARAIM (Baseline, no alert)

(f) SS ARAIM (Baseline, HUL)

Figure 6.13: Stanford diagrams for dual-frequency SPP using different approaches.

- **Satellite elevation mask** (15°). Signals with lower elevation are excluded;
- **Multipath LC criterion** ($3.89\,\sigma(\theta)$). Signals with greater LC values are excluded;
- **Ray tracing (RT) classification based on 3D city models.** Signals classified as NLOS or blocked are excluded;

Eight scenarios are defined regarding the applied functions, as detailed in Tab. 6.9.

It is noteworthy that Scenario 8 serves as a baseline, showcasing an ideal fault-free situation without utilizing any protection functions. Simulated measurements were generated based on

Table 6.9: Integrity and availability performance using different protection functions.

	Scenarios							
	1	2	3	4	5	6	7	8*
Functions applied:								
C/N0 mask		✓				✓	✓	
Elevation mask		✓				✓	✓	
MP LC criterion			✓		✓	✓	✓	
RT classification				✓	✓		✓	
Performence metric (FDE):								
HMI occurrence	6	2	6	1	1	2	1	0
MI occurrence	189	100	154	18	16	69	8	0
Availability [%]	92.1	89.8	91.3	72.9	72.5	88.6	70.6	96.6
Performence metric (FD-only):								
HMI occurrence	3	0	3	1	1	0	1	0
MI occurrence	171	87	142	13	13	59	8	0
Availability [%]	84.7	86.5	84.0	71.7	71.7	85.7	69.9	96.6

* Scenario 8 is a fault-free demonstration based on Monte Carlo simulations.

nominal error models and were used in the estimation instead of actual data. Specifically, interval-bounded error components were randomly generated using a uniform distribution, while the variance-described error components were handled by a Gaussian-distribution random generator.

Another notable aspect is the activation of full FDE functions in the interval approach. Once a detection occurs (i.e., the consistency check fails) at a particular epoch, the FE function is triggered. This function examines all potential fault modes to identify any existing faults and, if found, removes them from the estimation process (see Sec. 5.2.3). As a result, the outcomes may differ from those obtained using the FD function but without activating the FE function. For example, the protection functions used to generate the results in Fig. 6.13(b) are the same as those in Scenario 6, but the FE function is deactivated.

The evaluation results for the interval approach across eight scenarios, with and without activating the FE function, are summarized in Tab. 6.9 with selected scenarios illustrated using Stanford diagrams in Fig. 6.14.

The key observations from this study are as follows:

- The ideal simulated scenario (Scenario 8) demonstrates expected integrity monitoring performance, showing zero MI or HMI occurrences. However, the presence of remaining local threats could weaken its performance, as seen in other scenarios. To address this, either robust protection functions should be integrated or more precise information regarding the threat occurrence should be fed into the algorithms.
- Integrating more protection functions can reduce MI occurrences, e.g., the count drops significantly to 8 when all functions are employed, compared to 171 when no functions are used. However, this comes at the cost of availability performance, highlighting the trade-off between integrity and availability in choosing which functions to employ. The primary reason is that more active protection functions tend to exclude more satellites from processing, resulting in either insufficient satellite numbers or increased HPL values.

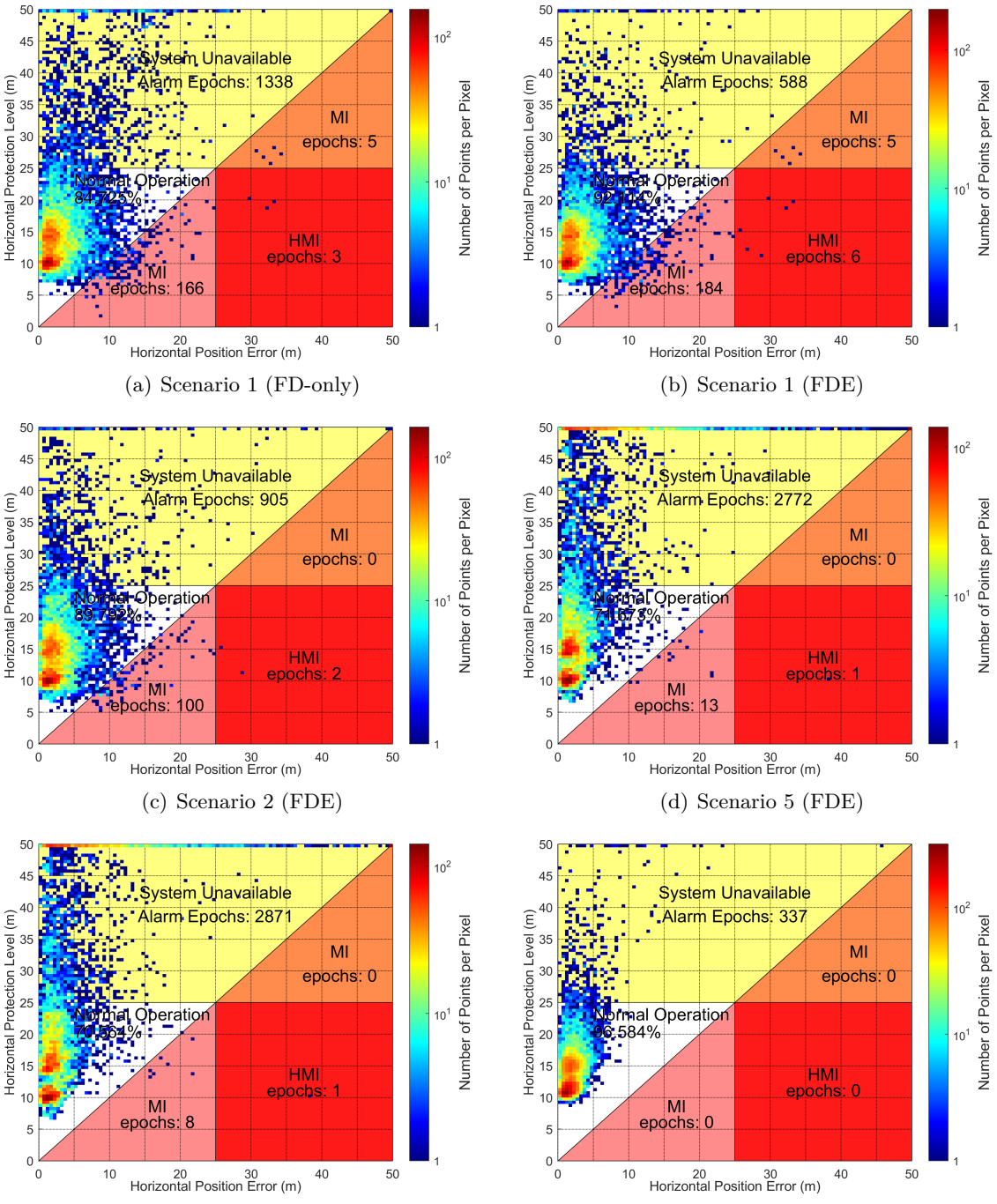


Figure 6.14: Stanford diagrams for the interval-based approach using different protection functions. Settings of the scenarios are referred to Tab. 6.9.

- The ray tracing classification proves particularly effective in identifying local threats. Scenarios that utilizes RT classification (4, 5,& 7) show fairly satisfactory integrity performance, with HMI occurrences meeting the IR requirement (1×10^{-4}) and relatively low MI occurrences.
- The impact of protection functions typically extends to FDE capability. The addition of the FE function improves availability performance by excluding detected faults. However, improper exclusion may cause LOI.

For example, in Scenario 1, there are 3 HMI and 171 MI events in the FD-only condition. When the FE function is activated, these numbers increase to 6 and 180, respectively, although this also results in more available epochs thanks to the exclusion of faulty measurements. This trend is also observed across other scenarios (2, 3, 4, 5 & 6). In contrast, the HMI and MI occurrences do not increase in Scenarios 7 & 8, and the availability in Scenario 7 improves slightly.

This difference is mainly attributed to the reliability (correctness) of the protection functions. If the protection functions yield unreliable results, the FE function, which relies on self-consistency checks, may be misled due to remaining faulty measurements and the absence of nominal ones. This shows a limitation of relying on GNSS alone navigation in challenging environments, necessitating the use of reliable external information or aiding sensors.

To further illustrate the impact, we choose a segment of 24 epochs from the experimental trajectory and observe their results in the position domain, as shown in Fig. 6.15(a) to Fig. 6.15(h). Detailed information regarding each epoch, including the number of satellites used in the estimation and the calculated HPL values, is provided in Fig. 6.15(i) and Fig. 6.15(j).

The improvement of integrity and degradation of availability is visible in, e.g., the third epoch (highlighted in Fig. 6.15(j)):

- In **Scenarios 1 and 2**, MI occurs as the reference point is not bounded by FSS, indicating loss of integrity;
- In **Scenarios 4, 5, and 7**, the system is unavailable as the obtained HPL values are beyond AL (dashed red line);
- **Scenarios 3, 6, and 8** illustrates no MI nor system unavailability, with the reference points bounded by FSS and HPL values remaining lower than AL.

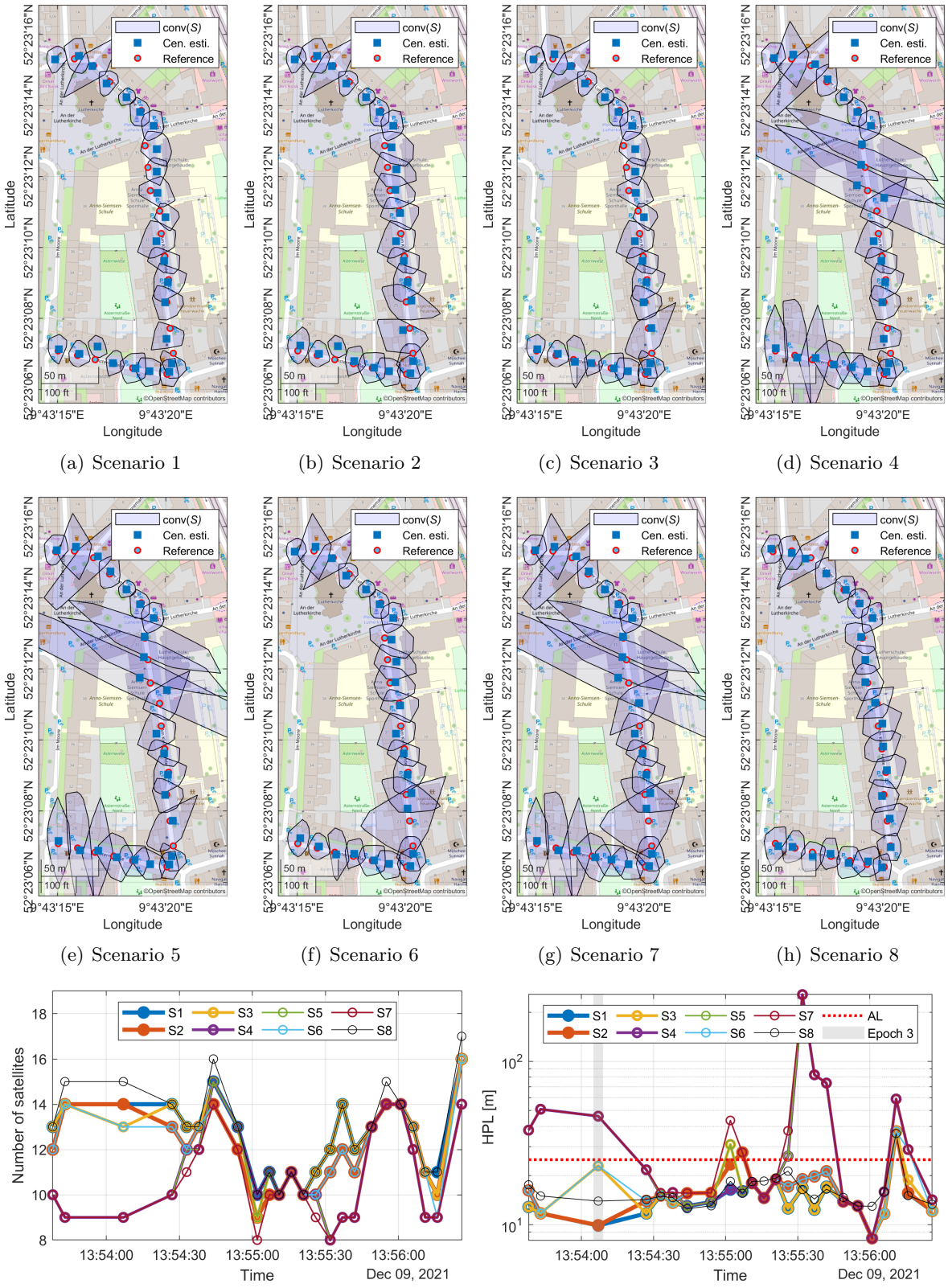
The experimental results highlight the importance of properly configuring protection functions against unknown local threats. On the one hand, unreliable protection functions can lead to incorrect fault identification, which may harm the system's nominal operations by excluding fault-free measurements. Moreover, the FDE capabilities of the integrity monitor are limited. Incorporating external information can help alleviate this issue, e.g., the RT classification based on 3D city models has demonstrated strong performance in this study.

On the other hand, GNSS observations can be largely affected by the challenging environments in urban areas. An excessive exclusion of faulty measurements may lead to an inadequate number of satellites for estimation and integrity monitoring, ultimately disrupting the navigation system. Therefore, the use of external sensors would be beneficial in addressing this challenge.

In conclusion, it is critical to seek a compromise between the integrity and availability of real-time integrity monitoring. It is highly recommended to explore external information or sensors to compensate for the limitations of GNSS in urban environments.

6.2.7 Discussion

This section has demonstrated the effectiveness of the interval-based integrity monitoring approach using two experimental datasets: one static dataset collected at a reference station and one kinematic dataset captured along a trajectory in a highly urbanized area.



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3 (d) Scenario 4

(e) Scenario 5 (f) Scenario 6 (g) Scenario 7 (h) Scenario 8

(i) Number of satellites used (eight scenarios) (j) HPL (eight scenarios)

Figure 6.15: Results of example epochs for the interval-based approach using different protection functions. Settings of the scenarios are referred to Tab. 6.9. The driving direction is anti-clockwise. The light-blue polygon represents the convex hull of the obtained non-convex feasible solution set (FSS).

The analysis for fault-free conditions confirms that the interval-based approach, by leveraging the interval-described nominal error bounding and potential fault removal, delivers significantly tighter position bounds compared to the classical stochastic approach. This indicates superior availability performance and highlights the advantages of this approach in applications such as infrastructure-based augmentation systems.

The interval approach is further evaluated in urban navigation scenarios with respect to LOI, in comparison with the legacy RB RAIM and the state-of-the-art baseline SS ARAIM algorithm. All three approaches succeed in ensuring the integrity for FD functions, with no MI occurrences in single-frequency SPP and no HMI occurrences in dual-frequency SPP. However, SS ARAIM failed in integrity assurance for alerted epochs, where excessive HMI events were observed.

Regarding system availability, RB RAIM consistently shows the worst performance due to its conservative calculation of position bounds, while the interval-based approach and SS ARAIM are more comparable. In single-frequency SPP scenarios, the interval approach produces notably lower HPL values in single constellation setups, reducing the values by up to 79.1% compared to RB RAIM and by 60.4% compared to SS ARAIM. With more constellations, however, it shows slightly increased HPL values, which is a reversed trend compared to the other approaches. For dual-frequency SPP, the interval-based approach achieves greater availability, outperforming RB RAIM by up to 25.5% and SS ARAIM by up to 12.5%, given the specified AL.

In urban environments with unknown local threats, reliable protection functions used for preprocessing are essential to maintain FDE capabilities. A crucial trade-off emerges: while excluding more potentially faulty measurements can enhance integrity, it may simultaneously degrade system availability. This trade-off underscores the importance of carefully balancing protection functions in GNSS-based navigation applications.

The analyses indicate that single-frequency pseudorange measurements alone are insufficient to meet demanding requirements. In contrast, dual-frequency pseudorange measurements demonstrate the potential for payment-critical or regulatory-critical applications (as outlined by EUSPA (2021)). Nevertheless, neither strategy fully achieves the required levels of integrity and availability for safety-critical applications due to inherent uncertainties in pseudorange observations.

To meet these stringent requirements, future research should explore the use of carrier-phase measurements in combination with the interval-based approach. Carrier-phase observations offer significantly higher precision, which could bridge the gap for safety-critical applications. Other recommendations from this chapter include enhancing protection functions that explore external information, improving availability through additional navigation sensors, and further refining FDE functions in future studies.

7 Conclusions and Outlook

In this dissertation, the key concepts, models, and methodologies for a comprehensive interval-based framework for autonomous integrity monitoring are presented and validated through both simulations and real-world experiments.

Unlike conventional stochastic approaches, this framework is grounded on interval-described uncertainty modeling. The starting point is the GNSS uncertainty component classification. Specifically, the remaining errors after applying standard correction methods to the GNSS measurements can be traced to various sources, yielding either stochasticity (describing the random variability of the measured values) or imprecision (caused by remaining systematics). Intervals offer an adequate description of uncertainty due to imprecision and can naturally be extended to probabilistic contexts, where it is interpreted as a confidence interval.

By leveraging interval mathematics and set theory, the interval-based approach outlined in this dissertation enables rigorous error bounding, deterministic uncertainty propagation, and fault detection and exclusion (FDE). This approach holds substantial relevance for a range of emerging integrity-focused applications beyond civil aviation, where the integrity concept was initially formalized. Its features are of great significance for a variety of safety-critical scenarios, such as Advanced Driver Assistance Systems (ADAS), railway signaling, and autonomous vehicle navigation. As highlighted in the report by EUSPA (2021), there is also a growing need for integrity monitoring among road users across various regulatory-critical and payment-critical GNSS applications.

Fig. 7.1 structures the road map toward an interval-based integrity monitoring framework. In line with this, five major methodological achievements in this dissertation are summarized as follows.

Achievement 1: Interval-described uncertainty modeling for GNSS pseudorange. Building on the foundation of distribution-free uncertainty modeling using the interval representation outlined in Schön (2003), this thesis introduces practical methods for characterizing uncertainty in GNSS pseudorange measurements. These approaches are tailored to quantify the uncertainty due to remaining errors after applying standard correction methods. In particular, three major error sources, including the residual tropospheric error, residual ionospheric error, and multipath error, have been evaluated and validated through real-world data. The key contributions include:

- The interval-based sensitivity analysis is refined using natural interval arithmetic, broadening its applicability to a wider range of problems, e.g., explicit nonlinear, discontinuous, or nonmonotonic models. It allows for individual estimation of the lower bound and upper bound and, hence, relaxes the limitation on symmetric intervals, potentially

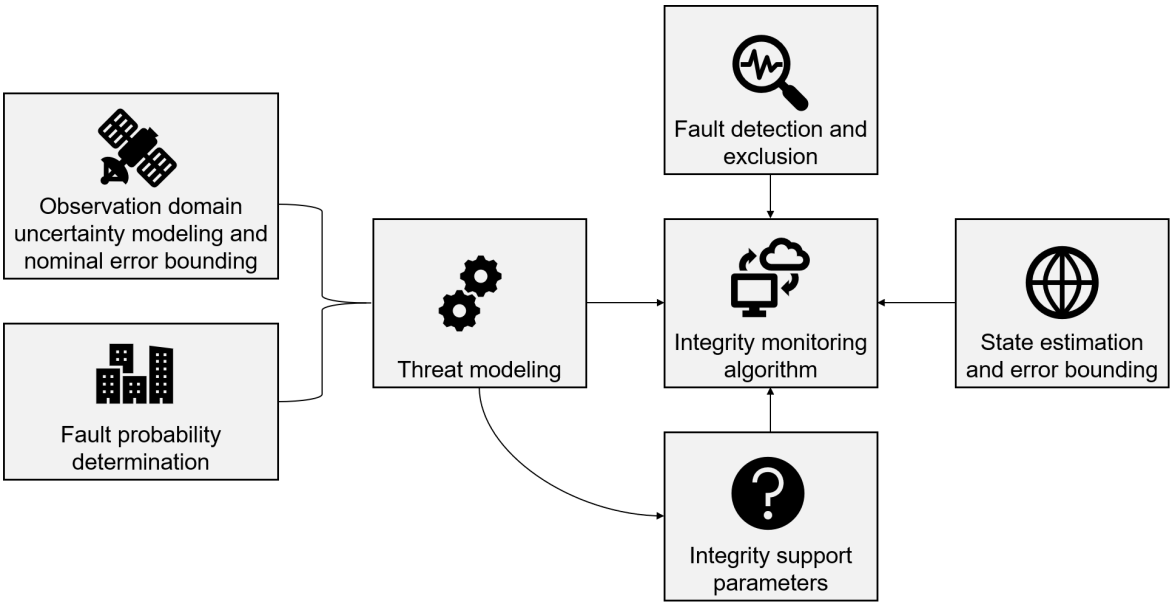


Figure 7.1: The road map toward an interval-based integrity monitoring framework developed in this thesis.

resulting in tighter bounds. This method has been used to address the residual tropospheric and ionospheric errors for the standard Saastamoinen model and GPS Klobuchar model, respectively.

- A multipath error envelope (MEE)-based method has been proposed and implemented to upper bound pseudorange multipath errors for binary phase shift keying (BPSK) and binary offset carrier (BOC) signals. Experimental evaluation for GPS, GLONASS, and Galileo signals demonstrates the effectiveness of this approach.

Achievement 2: Set-based state estimation and error bounding. The interval-described uncertainty models for GNSS state estimation problems have been explored in this dissertation, with an emphasis on error bounding. The classical least-squares estimator and the set estimator based on constraint satisfaction, examined in Schön (2003) and Dbouk (2021), are revisited, along with various error bounding approaches. Additionally, two novel point estimators are investigated that take advantage of the obtained feasible solution set (FSS), extending the scope of prior works:

- The optimal set-based central estimator produces a point solution as the Chebyshev center of the FSS, which is optimal in terms of minimal estimation error bounds.
- The set-constrained least-squares estimator integrates set constraints into the classical least-squares framework and, thereby combining the strength of both the set estimator and the least-squares estimator.

These estimators have been comprehensively analyzed and compared through theoretical and simulational analyses in terms of their mathematical properties, accuracy, and error bounding performance. Notably, in a two-dimensional example, the least-squares estimator demonstrated significant advantages in handling stochastic errors regarding both accuracy and error bounds, including cases of purely Gaussian-distributed errors and long-tailed distributed errors. The set-based central estimator is recommended for managing uniformly distributed errors (i.e., with known interval bounds) or hybrid errors that are dominated by interval-bounded components. This corresponds to application scenarios where there are significant remaining systematic errors. Benefiting from both the least-squares principle and set estima-

tor, the SCLS estimator offers robust performance, enabling the generation of a flexible and reliable point solution.

Achievement 3: Improvement of baseline ARAIM algorithms. A minor modification of the baseline ARAIM algorithms is suggested by this thesis, aimed at reducing the calculated fault-free HPL. This enhancement is achieved through the implementation of the interval extension of the least-squares estimator. The effectiveness of these improvements has been demonstrated via simulation analysis at both global and local scales.

Achievement 4: Set-based fault detection and exclusion. Following former investigations in Dbouk (2021), a detailed formulation of the set-based fault detection method is provided in this thesis, emphasizing its principle of detection that operates within the state domain. The set-based detector explores set operations and utilizes a set-emptiness check as the test criterion, eliminating the need to compute a scalar threshold. The mathematical properties of this method have been thoroughly analyzed for a simplified scalar estimation scenario using order statistics and compared to classical RAIM detectors, including RB and SS methods.

The detection interval characterizes the set-based detector and significantly influences both the false alert (FA) and missed detection (MD) probabilities, as well as its detection capability, which refers to its sensitivity to fault magnitudes. Assuming normally distributed errors, analytical expressions for these parameters have been explicitly derived. Additionally, a heuristic measure has been developed to qualify two-dimensional measurement geometry for set-based detection. This measure has been tested through MC simulation in comparison with the dilution of precision (DOP), which is commonly used in stochastic approaches. Furthermore, an innovative weighting scheme has been proposed to enhance detection capabilities.

Achievement 5: Interval-based integrity monitoring architecture. An interval-based integrity monitoring architecture has been developed in this thesis. It incorporates set-based fault detection and exclusion methods, along with the set-based central estimator, all grounded in interval-described uncertainty models. Leveraging the multiple-hypotheses framework from ARAIM, the architecture is capable of managing multiple simultaneous faults, protecting the navigation system from loss of integrity (LOI) with assured probability. This capability is particularly beneficial in urban navigation scenarios, where multiple simultaneous faults may arise due to challenging environments. A practical solution tailored for road transportation applications is presented, which aligns with the referenced Alert Limit (AL) framework by incorporating road geometry and varying levels of situational awareness.

The developed interval-based integrity monitoring approach has been evaluated in three ways:

- **Analytical evaluation for integrity risk (IR) bounds.** Two bounds have been defined and assessed for a benchmark example: (i) the NP bound, which is established using conditional No Protection (NP) probabilities, and (ii) the WCF bound, representing the maximum reachable HMI probability with corresponding worst-case fault (WCF). The former is computationally efficient but slightly conservative, while the latter, though tighter, requires prior computation of the WCF, which can be computationally expensive. Both bounds exhibit advantages over the residual-based (RB) and solution separation (SS) RAIM in cases of relatively large AL values or strict continuity risk (CR) requirements across various assumptions about the error distribution.
- **MC evaluation for integrity risk (IR) bounds**. Realistic scenarios have been analyzed using MC simulations, including a benchmark dual-constellation example and

24-hour changing geometries at an example location. Both the developed interval error bounds and the Gaussian error bounds derived from the guidance in the MOPS are involved in single and dual-frequency scenarios. In terms of system availability (lower HPL values are preferred), the interval-based approach outperforms the classical methods for single-frequency scenarios, although improvements are still needed for dual-frequency scenarios.

- **Experimental evaluation and validation for real-world urban navigation.** This evaluation uses data collected from urban environments, focusing on integrity and availability metrics. The threat model is established by adjusting the ARAIM model, generating nominal error bounds and a set of Integrity Support Parameters (ISP). Based on the threat modeling, the interval-based approach produces significantly tight position bounds for fault-free conditions, indicating superior availability performance for applications such as infrastructure-based augmentation systems. Compared to classical RB RAIM and the state-of-art baseline ARAIM algorithms, the developed approach consistently meets the IR requirements and shows higher availability in most scenarios. Furthermore, the reliability of protection functions against unknown local threats is emphasized as an important influence factor of the interval-based approach, as evidenced by the comparative analysis.

While the capabilities and advantages of the developed interval-based approach have been demonstrated, there are several aspects that offer promising opportunities for further development:

- The first aspect relates to the **interval-described GNSS uncertainty modeling**. In this thesis, example atmospheric correction models have been examined. Future research could explore additional application scenarios, such as the GPT series models for correcting the tropospheric delay, as well as the Galileo NeQuick model and Beidou BDGIM model for correcting the ionospheric delay. Another important direction is the treatment of residual atmospheric error and multipath error for carrier-phase measurements.
- The second aspect involves **extending state estimation beyond snapshot solutions** by exploring advanced filtering techniques to tackle dynamic issues. Key directions include deterministic uncertainty propagation through a filter structure and the development of associated error bounding approaches, which are fundamental to future integrity monitoring for multi-sensor fusion.
- To advance the continuous refinement of ARAIM for aviation users, future research, as the third aspect, could focus on **developing an advanced LSQ-derived estimator** that aims at minimizing position bounds and incorporates both variance-described and interval-described error components. This advancement would be beneficial not only in handling fault-free conditions but also in reducing the final HPL values.
- The fourth aspect concerns **improving set-based FDE methods**. Potential areas for investigation include:

  (i) exploring the mathematical characteristics of the set-based detector for higher-dimensional estimation problems,

  (ii) identifying meaningful geometry measures for high-dimensional measurement geometries, and

  (ii) developing robust weighting models, e.g., using environmental features or information from signal tracking.

- The last aspect is about the **integrity monitoring architecture**. In addition to the previous aspects that would ultimately benefit the integrity monitor, a number of recommendations can be made to facilitate real-world applications, including:

  (i) Realistic Integrity Support Parameter (ISP) configuration. Particularly, it is crucial to estimate and represent fault probabilities that account for local threats. Innovative strategies should be developed based on both satellite operational history and local environments. A promising solution is the *GNSS feature map* proposed by Ruwisch and Schön (2022) for urban navigation, which associates the parameter of interest with satellite elevations and driving locations.

  (ii) Integration of external information and aiding sensors. As highlighted in relevant sections, external information can be integrated into the architecture to support protection functions against unknown local threats. An example is the ray tracing (RT) classification based on 3D city models. Other GNSS measurements or additional sensors could help identify multipath-affected and NLOS signals, alleviating the reliance on self-consistency checks and enhancing FDE capabilities. For instance, using the GNSS Doppler shift measurements, as shown in Xu and Rife (2019), or various sensors such as odometers, fisheye cameras, LiDAR, and inertial sensors (Ahmad, 2013; Sánchez et al., 2016; Wen et al., 2019).

  (iii) Architecture for filter-based solutions. To address dynamic issues or capture time-correlated errors, the integrity monitoring architecture could be expanded by employing filtering techniques. Stochastic approaches, such as recursive ARAIM (Gallon, 2023), are currently under development. Future work could also investigate the multi-sensor application of the interval-based approach using filtering techniques.

A

Derivations for the benchmark problem using order statistics

This section presents relevant derivations for the benchmark problem in Sec. 5.1.3, including (i) the determination of feasible solution set (FSS) and (ii) probabilistic evaluation within the integrity context.

A.1 Determination of feasible solution set

Sec. 5.2.2 defines the measurement-level intervals, denoted by $[-\boldsymbol{\Delta}_s, \boldsymbol{\Delta}_s]$, and the associated bounding probability P_Δ. Four hypotheses, H_i for $i = 0, 1, 2, 3$, are established considering the single-measurement-fault assumptions:

Under H_0: The following inequality system:

$$\begin{cases} x \leq y_1 + \Delta_s, & -x \leq -y_1 + \Delta_s, \\ x \leq y_2 + \Delta_s, & -x \leq -y_2 + \Delta_s, \\ x \leq y_3 + \Delta_s, & -x \leq -y_3 + \Delta_s, \end{cases} \tag{A.1}$$

results in the fault-free polytope solution set:

$$\mathcal{P}_{s,0} = [\max(y_1, y_2, y_3) - \Delta_s, \min(y_1, y_2, y_3) + \Delta_s]. \tag{A.2}$$

Under H_1: The measurement y_1 is assumed to be faulty. Since there is no information available regarding the fault magnitude, nor should any assumption be made, we inflate its interval bounds to infinity. As a result, the following inequalities are obtained:

$$\begin{cases} x \leq y_1 + \infty, & -x \leq -y_2 + \infty, \\ x \leq y_2 + \Delta_s, & -x \leq -y_2 + \Delta_s, \\ x \leq y_3 + \Delta_s, & -x \leq -y_3 + \Delta_s, \end{cases} \tag{A.3}$$

resulting in a new polytope solution set:

$$\mathcal{P}_{s,1} = [\max(y_2, y_3) - \Delta_s, \min(y_2, y_3) + \Delta_s], \tag{A.4}$$

which tolerates the potential faulty measurement y_1. In this regard, $\mathcal{P}_{s,1}$ is interpreted as a polytope solution set that tolerates potential fault in measurement y_1 (termed fault-tolerant polytope solution set hereafter). For simplicity of expressions, the inequalities with tolerant measurements (e.g., the first line in Eq. A.3) can be neglected without altering the calculation result.

Under H_2: Similar to H_1, a fault-tolerant polytope solution set is obtained when y_2 is tolerated (bounded by an infinitely-wide interval):

$$\mathcal{P}_{s,2} = [\max(y_1, y_3) - \Delta_s, \min(y_1, y_3) + \Delta_s]. \tag{A.5}$$

Under H_3: The corresponding fault-tolerant polytope set solution is expressed as

$$\mathcal{P}_{s,3} = [\max(y_1, y_2) - \Delta_s, \min(y_1, y_2) + \Delta_s]. \tag{A.6}$$

Subsequently, the feasible solution set (FSS, $\mathcal{S}$) is the union of all the above polytopes:

$$\begin{aligned}
\mathcal{S} =& \mathcal{P}_{s,0} \cup \mathcal{P}_{s,1} \cup \mathcal{P}_{s,2} \cup \mathcal{P}_{s,3} \\
=& [\max(y_1, y_2, y_3) - \Delta_s, \min(y_1, y_2, y_3) + \Delta_s] \cup [\max(y_2, y_3) - \Delta_s, \min(y_2, y_3) + \Delta_s] \\
& \cup [\max(y_1, y_3) - \Delta_s, \min(y_1, y_3) + \Delta_s] \cup [\max(y_1, y_2) - \Delta_s, \min(y_1, y_2) + \Delta_s] \\
=& [\min(\max(y_1, y_2, y_3), \max(y_2, y_3), \max(y_1, y_3), \max(y_1, y_2)) - \Delta_s, \\
& \max(\min(y_1, y_2, y_3), \min(y_2, y_3), \min(y_1, y_3), \min(y_1, y_2)) + \Delta_s] \\
=& [Y_{(2:3)} - \Delta_s, Y_{(2:3)} + \Delta_s],
\end{aligned} \tag{A.7}$$

where $Y_{(2:3)}$ denotes the second order statistic out of three samples (the order statistics is introduced in Eq. A.14), i.e., the median of three samples. Eq. A.7 implies the FSS to be an interval around $Y_{(2:3)}$ with radius Δ_s. Hence, the optimal point estimator is the median estimator $\hat{\mathbf{x}}_s = Y_{(2:3)}$, and the scalar measure of $\mathcal{S}$, indicating the worst-case position error, is $\text{meas}(\mathcal{S}) = \Delta_s$.

A.2 Probabilistic evaluation

A.2.1 Introduction

This section evaluates the relevant probabilities for the benchmark problem. The PDF and CDF of measurements are denoted by $f_i(y)$ and $F_i(y)$ for $i = 1, 2, 3$. Under the fault-free hypothesis (H_0, with the fault vector $\mathbf{f} = [0, 0, 0]^T$), they follow normal distributions:

$$\begin{gathered}
f_1(y) = f_2(y) = f_3(y) = f_N(y),\ F_1(y) = F_2(y) = F_3(y) = F_N(y) \\
\Rightarrow \mathbf{y} \sim \mathcal{N}(\mathbf{0}_{3\times1}, \mathbf{I}_{3\times3}),
\end{gathered} \tag{A.8}$$

where $f(y)$ and $F(y)$ denote the PDF and CDF of a standard normal distribution.

Under a faulty hypothesis (H_i for $i = 1, 2, 3$), the faulty measurement follows a non-zero mean normal distribution (PDF denoted by $f_N(y, f_i)$ and CDF denoted by $F_N(y, f_i)$):

$$\begin{gathered}
f_i(y) = f_N(y, f_i),\ F_i(y) = F_N(y, f_i) \\
\Rightarrow \mathbf{y} \sim \mathcal{N}(\mathbf{f}, \mathbf{I}_{3\times3}),
\end{gathered} \tag{A.9}$$

where $\mathbf{f}$ is defined in Eq. 5.13, dependent on the fault mode.

A.2.2 False alert probability

Without loss of generality, the joint PDF of two order statistics $Y_{(r:n)}$ and $Y_{(s:n)}$ ($1 \le r < s \le n$) is denoted by (see David and Nagaraja (2003), Page 11)

$$\begin{aligned}
f_{(r:n)(s:n)}(y_1, y_2) = & \frac{n!}{(r-1)!(s-r-1)!(n-s)!} \cdot \\
& F_N^{r-1}(y_1)\, f_N(y_1)\, [F_N(y_2) - F_N(y_1)]^{s-r-1}\, f_N(y_2)\, [1 - F_N(y_2)]^{n-s}.
\end{aligned} \tag{A.10}$$

The joint CDF may be obtained by integration of Eq. A.10, or by a direct argument in the discrete case for $y_1 < y_2$:

$$\begin{aligned} F_{(r:n)(s:n)}(y_1, y_2) =& P\left(\text{at least } r \; Y_i \leq y_1, \text{ at least } s \; Y_i \leq y_2\right) \\ =& \sum_{j=s}^{n} \sum_{i=r}^{j} P\left(\text{exactly } i \; Y_i \leq y_1, \text{ exactly } j \; Y_i \leq y_2\right) \\ =& \sum_{j=s}^{n} \sum_{i=r}^{j} \frac{n!}{i!(j-i)!(n-j)!} F_N^i(y_1) \left[F_N(y_2) - F_N(y_1)\right]^{j-i} \\ & \cdot \left[1 - F_N(y_2)\right]^{n-j}. \end{aligned} \tag{A.11}$$

For $y_1 > y_2$, Eq. A.11 can be simplified as

$$F_{(r:n)(s:n)}(y_1, y_2) = F_{(s:n)}(y_2). \tag{A.12}$$

Defining new terms $w \coloneqq y_2 - y_1, y \coloneqq y_1$, Eq. A.10 may be re-arranged concerning the range of the two order statistics $W_{rs} \coloneqq Y_{(s)} - Y_{(r)}$ by transforming from y_1, y_2 to y, w. Subsequently, the PDF of W_{rs} is obtained by integrating over y for Eq. A.10:

$$\begin{aligned} f_{W_{rs}(w)} =& \int_{-\infty}^{\infty} f_{(1:n)(n:n)}(y, y+w) \, \mathrm{d}y \\ =& \frac{n!}{i!(j-i)!(n-j)!} F^i(y) \cdot \\ & \int_{-\infty}^{\infty} F^{r-1}(y) \, f(y) \left[F(y_1 + w) - F(y)\right]^{s-r-1} f(y+w) \left[1 - F(y+w)\right]^{n-s} \mathrm{d}x. \end{aligned} \tag{A.13}$$

The computation of FA probability utilizes the definition of "test statistic" W (sample range) in Eq. 5.19. Its PDF can be obtained as a special case $r = 1$, $s = n$ of Eq. A.13, and its CDF can be obtained by integrating the PDF over $w > 0$ and $y \in [-\infty, \infty]$, resulting in formulations in Eq. 5.20. Consequently,

$$P(\text{FA}) = \int_{2\,\Delta_s}^{\infty} f_W(w) \mathrm{d}w = 1 - F_W(2\,\Delta_s). \tag{A.14}$$

Notably, Eq. A.14 and Eq. 5.21 evaluate the FA probability using the same method, with the only difference in the notation of intervals involved (Δ_d in Eq. 5.21 and Δ_s in Eq. A.14).

A.2.3 No detection probability

Under H_0: The conditional ND probability is the complementary probability to the FA probability:

$$P(\text{ND} \mid H_0) = 1 - P(\text{FA}) = F_W(2\Delta_s), \tag{A.15}$$

where $F_W(w)$ is defined in Eq. 5.20.

Under H_i: The order statistics used in the fault-free hypothesis (H_0) is no longer applicable, because a combinatorial problem is formed under H_i, i.e., three order statistics with one candidate faulty measurement. In this regard, the *independent identically distributed (iid)* assumption must be discarded, and *inid* order statistics (i.e., *independent nonidentically distributed* variables) should be explored. The derivation is as follows.

The joint PDF $f_{(1:3)(3:3)}(y, y+w)$ of the first and third order statistics can be expressed explicitly as

$$\begin{aligned} f_{(1:3)(3:3)}(y, y+w) = \\ & f_1(y)\,[F_2(y+w) - F_2(y)]\,f_3(y+w) + f_1(y)\,[F_3(y+w) - F_3(y)]\,f_2(y+w) + \\ & f_2(y)\,[F_1(y+w) - F_1(y)]\,f_3(y+w) + f_2(y)\,[F_3(y+w) - F_3(y)]\,f_1(y+w) + \\ & f_3(y)\,[F_1(y+w) - F_1(y)]\,f_3(y+w) + f_3(y)\,[F_2(y+w) - F_2(y)]\,f_1(y+w). \end{aligned} \tag{A.16}$$

Generalizations for n measurements are

$$\begin{aligned} f_{(1:n)(n:n)}(y, y+w) = \sum_{k=1}^{2^n} & f_{\mathrm{perm}_k(1)}(y) \cdot [F_{\mathrm{perm}_k(2)}(y+w) - F_{\mathrm{perm}_i(2)}(y)] \cdot ... \\ & \cdot [F_{\mathrm{perm}_k(n-1)}(y+w) - F_{\mathrm{perm}(n-1)}(y)] \cdot f_{\mathrm{perm}_k(n)}(y+w), \end{aligned} \tag{A.17}$$

where $\mathrm{perm}_k(j)$ for $k \in [1, 2^n]$ and $j \in [1, n]$ represents the j-th element in the k-th possible permutation for n measurements; $f_{\mathrm{perm}_k(1)}(y) = f_N(y)$ (also $F_{\mathrm{perm}_k(1)}(y) = F_N(y)$) when $k = i$ and $f_{\mathrm{perm}_k(1)}(y) = f_N(y, f_i)$ (also $F_{\mathrm{perm}_k(1)}(y) = F_N(y, f_i)$) when $k \neq i$.

The CDF of W can be derived from integration over w and y:

$$\begin{aligned} F_W(w) &= \int_0^\infty \int_{-\infty}^\infty f_{(1:n)(n:n)}(y, y+w)\,\mathrm{d}y\,\mathrm{d}w \\ &= \int_{-\infty}^\infty \sum_{k=1}^{2^n} f_{\mathrm{perm}_k(1)}(y) \cdot \int_0^\infty [F_{\mathrm{perm}_k(2)}(y+w) - F_{\mathrm{perm}_k(2)}(y)] \cdot ... \\ &\quad \cdot [F_{\mathrm{perm}_k(n)}(y+w) - F_{\mathrm{perm}(n-1)}(y)] \cdot f_{\mathrm{perm}_k(n)}(y+w)\,\mathrm{d}w\,\mathrm{d}y. \end{aligned} \tag{A.18}$$

Undoubtly, Eq. 5.20 is a special case (*iid* case) of Eq. A.17 and Eq. A.18.

Subsequently, the conditional ND probability is determined by

$$P(\mathrm{ND} \mid H_i) = F_W(2\Delta_s), \tag{A.19}$$

with $F_W(w)$ defined in Eq. A.18. Accordingly, Fig. 5.4(b) illustrates the results of Eq. A.18 as Δ_s and f_i vary when $n = 3$.

A.2.4 Probability of hazardous misleading information

Under H_0: The conditional HMI probability can be analytically assessed by

$$\begin{aligned} P(\mathrm{HMI} \mid H_0) &= P\left(|\varepsilon_0| > \ell \wedge \mathcal{P}_{s,0} \neq \emptyset \mid \mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{3\times3})\right) \\ &= P\left(Y_{(3:3)} - Y_{(1:3)} < 2\Delta_s \wedge \Big| Y_{(2:3)}| > \Delta_s \mid \mathbf{y} \sim \mathcal{N}(\mathbf{0}_{n\times1}, \mathbf{I}_{3\times3})\right). \end{aligned} \tag{A.20}$$

David and Nagaraja (2003) (Page 12) expresses the joint PDF of all n *iid* order statistics as

$$f_{(1:n),...,(n:n)}(y_1, ..., y_n) = n! f_N(y_1) \cdot ... \cdot f_N(y_n) \quad \text{with } y_1 \leq ... \leq y_n. \tag{A.21}$$

Defining the following variables:

$$w = Y_{(3:3)} - Y_{(1:3)} \in [0, \infty), \quad r = Y_{(2:3)} - Y_{(1:3)} \in [0, w], \quad y = Y_{(2:3)}, \tag{A.22}$$

the joint PDF of three observations can be formulated as

$$f_{(1:3),(2:3),(3:3)}(y_1, y_2, y_3) = f_{y,w,r}(y, w, r) = 3! f_N(y - r)\, f_N(y)\, f_N(y - r + w). \tag{A.23}$$

Eq. A.20 reveals $w \geq 0$, $w < 2\Delta_s$ and $y > \Delta_s$, and hence, the conditional HMI probability is calculated by integration over the three variables:

$$\begin{aligned} &P(\mathrm{HMI} \,|\, H_0) \\ =&\int_{-\infty}^{-\Delta_s} \int_0^{\Delta_s} \int_0^{w} f_{y,w,r}(y,w,r)\mathrm{d}r\,\mathrm{d}w\,\mathrm{d}y + \int_{\Delta_s}^{\infty} \int_0^{\Delta_s} \int_0^{w} f_{y,w,r}(y,w,r)\mathrm{d}r\,\mathrm{d}w\,\mathrm{d}y \\ =&3!\left(\int_{-\infty}^{-\Delta_s} f_N(y) \int_0^{\Delta_s} \int_0^{w} f_N(y-r) f_N(y-r+w)\mathrm{d}r\,\mathrm{d}w\,\mathrm{d}y \right) \\ &+ \int_{\Delta_s}^{\infty} f_N(y) \int_0^{\Delta_s} \int_0^{w} f_N(y-r) f_N(y-r+w)\mathrm{d}r\,\mathrm{d}w\,\mathrm{d}y \Bigg). \end{aligned} \tag{A.24}$$

Under H_i: The joint PDF of three observations is formulated by considering various permutations:

$$\begin{aligned} f^{(i)}_{y,w,r}(y,w,r) = 2\, f_N(y,f_i) f_N(y-r) f_N(y-r+w) + 2\, f_N(y) f_N(y-r,f_i) f_N(y-r+w) + \\ 2\, f_N(y) f_N(y-r) f_N(y-r+w,f_i), \end{aligned} \tag{A.25}$$

where the fault magnitude f_i is arbitrary and the superscript (i) is associate with H_i, indicating the *inid* situation different from Eq. A.21. The permutations are exhaustively listed to account for all possible orders of measurement magnitudes.

The conditional HMI probability under H_i is then given by

$$\begin{aligned} &P(\mathrm{HMI} \,|\, H_i) \\ =&P(|\varepsilon_0| > \ell \wedge \mathcal{P}_{s,0} \neq \emptyset \,|\, \mathbf{y} \sim \mathcal{N}(\mathbf{f}_i, \mathbf{I}_{3\times 3})) \\ =&P(Y_{(3:3)} - Y_{(1:3)} < 2\Delta_s \wedge |Y_{(2:3)}| > \Delta_s \,|\, \mathbf{y} \sim \mathcal{N}(\mathbf{f}_i, \mathbf{I}_{3\times 3})) \\ =&\int_{-\infty}^{-\Delta_s} \int_0^{2\,\Delta_s} \int_0^{w} f^{(i)}_{y,w,r}(y,w,r)\mathrm{d}r\,\mathrm{d}w\,\mathrm{d}y + \int_{\Delta_s}^{\infty} \int_0^{2\,\Delta_s} \int_0^{w} f^{(i)}_{y,w,r}(y,w,r)\mathrm{d}r\,\mathrm{d}w\,\mathrm{d}y. \end{aligned} \tag{A.26}$$

In the benchmark example, Eq. A.26 is consistent across all faulty hypotheses.

The remaining question is to determine the value of f_i to be evaluated, which is arbitrary by definition. Similar to the evaluation method for RB RAIM (cf. Eq. 2.65), the HMI probability for the interval approach can also be evaluated against the wort-case fault magnitude $f_{i,\mathrm{worst}}$:

$$\begin{aligned} P(\mathrm{HMI} \mid H_i) \leq& P(\mathrm{HMI} \mid \mathbf{f}_{i,\mathrm{worst}}) \\ =& P(|\varepsilon_0| > \ell \wedge \mathcal{P}_{s,0} \neq \emptyset \mid \mathbf{y} \sim \mathcal{N}(\mathbf{f}_{i,\mathrm{worst}}, \mathbf{I}_{3\times 3})), \end{aligned} \tag{A.27}$$

where $f_{i,\mathrm{worst}}$ is the i-th element of $\mathbf{f}_{i,\mathrm{worst}}$ and the other two elements are zeros. $P(\mathrm{HMI} \mid H_i)$ is maximized for H_i in the case of $\mathbf{f}_{i,\mathrm{worst}}$. Replacing f_i with $f_{i,\mathrm{worst}}$ in Eq. A.25 and using methods such as line search, $P(\mathrm{HMI} \mid \mathbf{f}_{i,\mathrm{worst}})$ and $f_{i,\mathrm{worst}}$ can be computed jointly.

B

Formulas for Horizontal Protection Level

Various formulas for calculating Horizontal Protection Level (HPL)/Horizontal Uncertainty Level (HUL) can be found in the literature. In addition to the methods introduced in Sec. 2.4, this section reports three approaches that are used in the experiments in Sec. 6.2.

MOPS SBAS-style HPL. The MOPS provides a formula for computing the SBAS HPL (RTCA/DO-229D, 2006):

$$\mathrm{HPL} = k_H \cdot d_{major}, \tag{B.1}$$

where

- d_{major} may be interpreted as the semi-major axis of the confidence ellipse for the least-squares estimation.
- k_H is a coefficient, which MOPS specifies 6.18 or 6.0 for different aircraft operations based on the required IR allocation. Its value is determined based on knowledge of position uncertainty (Eq. 4.7) and specific IR requirements:

$$1 - I_{REQ,H} = \int_{-\infty}^{k_H^2} \chi^2(2,0)\,\mathrm{d}x, \tag{B.2}$$

  where $\chi^2(2,0)$ represents a chi-square distribution with two degrees of freedom and $I_{REQ,H}$ is the IR requirement allocated to horizontal components. Notably, Eq. B.2 is slightly different from the handling of MOPS but is consistent with the error bounding as discussed in Sec. 4.2.

Legacy RB RAIM HUL. Different from the HPL in RB RAIM, the Horizontal Uncertainty Level (HUL) is a position error bound that meets the integrity requirement but is directly dependent on the measurement residuals (WG-C ARAIM TSG, 2019).

It is known from Eq. 2.44 that the parity vector can be interpreted as a function of the nominal error ($\mathbf{e}$) and the fault ($\mathbf{f}$). As a result, its L_2-norm, equal to q_{RB}^2, cannot fully capture the influence of $\mathbf{f}$ and must be compensated for to upper bound the position error. To do so, an upper bound is established as follows:

$$\mathrm{HUL} = \mathrm{HSLOPE}_{\max} \cdot (q_{RB}^2 + \Delta p) + k_H \cdot \sigma_{\mathrm{bias}}, \tag{B.3}$$

where Δp is added to compensate for the impact of $\mathbf{e}$. Given p_{bias} and T_{RB}, the Δp term may be calculated by

$$\Delta p = p_{\mathrm{bias}} - T_{RB}, \tag{B.4}$$

which can be understood as the impact of the noise term on the test statistic q_{RB} such that it is neither below the detection threshold ($q_{RB} < T_{RB}$, i.e., missed detection), nor approaching p_{bias}.

Furthermore, an empirically determined factor 1/2 may be applied to Δp, as proposed by Young et al. (1996), which can reduce the conservatism and has been verified by Monte Carlo simulations. Subsequently, the final formulation of HUL reads:

$$\text{HUL} = \text{HSLOPE}_{\max} \cdot \left(q_{RB}^2 + \frac{1}{2} \left(p_{\text{bias}} - T_{RB} \right) \right) + k_H \cdot \sigma_{\text{bias}}. \tag{B.5}$$

Baseline ARAIM HUL. Similar to HPL, the HUL in ARAIM is calculated using the two horizontal components (WG-C ARAIM TSG, 2019):

$$\text{HUL} = \sqrt{\text{HUL}_{(1)}^2 + \text{HUL}_{(2)}^2}, \tag{B.6}$$

where

$$\text{HUL}_{(m)} = \max_{k=0,\ldots,h} \left| \hat{x}_{i,(m)} - \hat{x}_{0,(m)} \right| + k_{i,(m)} \sigma_{i,(m)}, \tag{B.7}$$

and $k_{i,(m)}$ is known from PL computation:

$$k_{i,(m)} = \frac{\text{HPL}_{(m)} - T_{SS,i,(m)} - |\mathbf{k}_{(m),i}| \cdot \mathbf{b}_{nom}}{\sigma_{i,(m)}}. \tag{B.8}$$

Bibliography

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., et al. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information fusion*, 76:243–297.

Adjrad, M. and Groves, P. D. (2018). Intelligent urban positioning: Integration of shadow matching with 3D-mapping-aided GNSS ranging. *The Journal of Navigation*, 71(1):1–20.

Ahmad, K. (2013). Reliability of GNSS measurements via pseudorange prediction using an odometer for robust land-vehicle navigation. In: *Proceedings of the 26th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2013)*, pp. 1299–1307.

Alamo, T., Bravo, J. M., and Camacho, E. F. (2005). Guaranteed state estimation by zonotopes. *Automatica*, 41(6):1035–1043.

Alquist, J. L. and Baumeister, R. F. (2023). Dealing with uncertain situations. *The Journal of Positive Psychology*, pp. 1–24.

Althoff, M. (2010). *Reachability analysis and its application to the safety assessment of autonomous cars*. Diss., Technische Universität München.

Amir, D. (1978). Chebyshev centers and uniform convexity. *Pacific Journal of Mathematics*, 77(1):1–6.

Angus, J. E. (2006). RAIM with multiple faults. *Navigation*, 53(4):249–257.

Askne, J. and Nordius, H. (1987). Estimation of tropospheric delay for microwaves from surface weather data. *Radio Science*, 22(03):379–386.

Axmann, J., Moftizadeh, R., Su, J., Tennstedt, B., Zou, Q., Yuan, Y., Ernst, D., Alkhatib, H., Brenner, C., and Schön, S. (2023). LUCOOP: Leibniz University cooperative perception and urban navigation dataset. In: *2023 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1–8. IEEE.

Barnes, E. R. and Moretti, A. C. (2005). Some results on centers of polytopes. *Optimization Methods and Software*, 20(1):9–24.

Beer, M., Kougioumtzoglou, I. A., and Patelli, E. (2014). *Maintenance and safety of aging infrastructure*, Chap. Emerging concepts and approaches for efficient and realistic uncertainty quantification, pp. 121–154. CRC.

Blanch, J., Ene, A., Walter, T., and Enge, P. (2007). An optimized multiple hypothesis RAIM algorithm for vertical guidance. In: *Proceedings of the 20th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2007)*, pp. 2924–2933, Fort Worth, TX, USA.

Blanch, J., Walker, T., Enge, P., Lee, Y., Pervan, B., Rippl, M., Spletter, A., and Kropp, V. (2015a). Baseline Advanced RAIM user algorithm and possible improvements. *IEEE Transactions on Aerospace and Electronic Systems*, 51(1):713–732.

Blanch, J. and Walter, T. (2021). A fault detection and exclusion estimator designed for integrity. In: *Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*, pp. 1672–1686.

Blanch, J. and Walter, T. (2022). Design of estimators with integrity in the presence of error model uncertainty. In: *Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022)*, pp. 1467–1480.

Blanch, J., Walter, T., Berz, G., Burns, J., Clark, B., Joerger, M., Mabilleau, M., Martini, I., Milner, C., Pervan, B., et al. (2019a). Development of Advanced RAIM minimum operational performance standards. In: *Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019)*, pp. 1381–1391, Miami, Florida.

Blanch, J., Walter, T., and Enge, P. (2010). RAIM with optimal integrity and continuity allocations under multiple failures. *IEEE Transactions on Aerospace and Electronic Systems*, 46(3):1235–1247.

Blanch, J., Walter, T., and Enge, P. (2013). Optimal positioning for Advanced RAIM. *Navigation: Journal of the Institute of Navigation*, 60(4):279–289.

Blanch, J., Walter, T., and Enge, P. (2015b). Fast multiple fault exclusion with a large number of measurements. In: *Proceedings of the 2015 International Technical Meeting of The Institute of Navigation*, pp. 696–701.

Blanch, J., Walter, T., and Enge, P. (2018). Fixed subset selection to reduce Advanced RAIM complexity. In: *Proceedings of the 2018 International Technical Meeting of The Institute of Navigation*, pp. 88–98.

Blanch, J., Walter, T., and Enge, P. (2019b). Gaussian bounds of sample distributions for integrity analysis. *IEEE Transactions on Aerospace and Electronic Systems*, 55(4):1806–1815.

Blesa, J., Puig, V., Romera, J., and Saludes, J. (2011). Fault diagnosis of wind turbines using a set-membership approach. *IFAC Proceedings Volumes*, 44(1):8316–8321.

Blesa, J., Puig, V., and Saludes, J. (2012). Robust fault detection using polytope-based set-membership consistency test. *IET Control Theory & Applications*, 6(12):1767–1777.

Böhm, J., Möller, G., Schindelegger, M., Pain, G., and Weber, R. (2015). Development of an improved empirical model for slant delays in the troposphere (GPT2w). *GPS Solutions*, 19(3):433–441.

Boyd, S. P. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Braasch, M. S. (1997). Autocorrelation sidelobe considerations in the characterization of multipath errors. *IEEE Transactions on Aerospace and Electronic Systems*, 33(1):290–295.

Braasch, M. S. (2017). *Springer Handbook of Global Navigation Satellite Systems*, Chap. Multipath, pp. 443–468. In: Teunissen and Montenbruck (2017).

Brenner, M. (1995). Integrated GPS/inertial fault detection availability. In: *Proceedings of the 8th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1995)*, Palm Springs, CA, USA.

Brown, R. G. (1992). A baseline GPS RAIM scheme and a note on the equivalence of three RAIM methods. *Navigation*, 39(3):301–316.

Brown, R. G. and Chin, G. Y. (1997). GPS RAIM: Calculation of threshold and protection radius using chi-square methods-a geometric approach. *Global Positioning System: Papers Published in NAVIGATION*, V(10):155–179.

Caizzone, S., Circiu, M.-S., Elmarissi, W., Enneking, C., Rippl, M., and Sgammini, M. (2022). The role of antennas on GNSS pseudorange and multipath errors and their impact on DFMC multipath models for avionics. *NAVIGATION: Journal of the Institute of Navigation*, 69(3).

Campos, F., Neves, A., and de Souza, F. M. C. (2007). Decision making under subjective uncertainty. In: *2007 IEEE Symposium on Computational Intelligence in Multi-Criteria Decision-Making*, pp. 85–90. IEEE.

Casini, M. (2002). *Set-membership estimation: an advanced tool for system identification*. Diss., Univesita Degli Studi Di Siena.

Chen, G., Wang, J., and Shieh, L. S. (1997). Interval Kalman filtering. *IEEE Transactions on Aerospace and electronic Systems*, 33(1):250–259.

Chisci, L., Garulli, A., and Zappa, G. (1996). Recursive state bounding by parallelotopes. *Automatica*, 32(7):1049–1055.

Cimino, M. G., Lazzerini, B., Marcelloni, F., and Pedrycz, W. (2014). Genetic interval neural networks for granular data regression. *Information Sciences*, 257:313–330.

Collins, J. P., Langley, R., et al. (1999). Nominal and extreme error performance of the UNB3 tropospheric delay model. Final contract report for Nav Canada Satellite Navigation Program Office, Geodetic Research Laboratory, Department of Geodesy and Geomatics Engineering Technical Report No. 204, University of New Brunswick, Fredericton, New Brunswick, Canada.

Collins, J. P. and Langley, R. B. (1997). A Tropospheric Delay Model for the user of the Wide Area Augmentation System. Final contract report prepared for Nav Canada Satellite Navigation Program Office, Geodetic Research Laboratory, Department of Geodesy and Geomatics Engineering Technical Report No. 187, University of New Brunswick, Fredericton, New Brunswick, Canada.

Combastel, C. (2003). A state bounding observer based on zonotopes. In: *2003 European control conference (ECC)*, pp. 2589–2594. IEEE.

Combastel, C. (2016). An Extended Zonotopic and Gaussian Kalman Filter (EZGKF) merging set-membership and stochastic paradigms: Toward non-linear filtering and fault detection. *Annual Reviews in Control*, 42:232–243.

Conrad, J., Botner, O., Hallgren, A., and de Los Heros, C. P. (2003). Including systematic uncertainties in confidence interval construction for Poisson statistics. *Physical Review D*, 67(1):012002.

Crespillo, O. G., Gonzalez, R. O., and Caizzone, S. (2024). Airborne time-correlated GNSS multipath error modeling of carrier-phase smoothed code. In: *Proceedings of the 2024 International Technical Meeting of The Institute of Navigation*, pp. 301–306.

Crespillo, O. G., Grosch, A., Adjroloh, P. H., Zhu, C., Capua, R., Frittella, F., and Kutik, O. (2022). Multisensor localization architecture for high-accuracy and high-integrity land-based applications. In: *Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022)*, pp. 1873–1889.

Dach, R., Brockmann, E., Schaer, S., Beutler, G., Meindl, M., Prange, L., Bock, H., Jäggi, A., and Ostini, L. (2009). GNSS processing at CODE: status report. *Journal of Geodesy*, 83:353–365.

Dach, R., Schaer, S., Arnold, D., Brockmann, E., Kalarus, M. S., Prange, L., Stebler, P., and Jäggi, A. (2023). CODE final product series for the IGS [Dataset]. https://boris.unibe.ch/id/eprint/185744. Accessed: 2024-01-10.

David, H. A. and Nagaraja, H. N. (2003). *Order statistics*. John Wiley & Sons, Hoboken, USA, 3 edition.

Davis, J., Herring, T., Shapiro, I., Rogers, A., and Elgered, G. (1985). Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length. *Radio science*, 20(6):1593–1607.

Dbouk, H. (2021). *Alternative Integrity Measures Based on Interval Analysis and Set Theory*. Diss., Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover.

Dbouk, H. and Schön, S. (2018). Comparison of different bounding methods for providing GPS integrity information. In: *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pp. 355–366. IEEE.

Dbouk, H. and Schön, S. (2019). Reliability and integrity measures of GPS positioning via geometrical constraints. In: *Proceedings of the 2019 International Technical Meeting of The Institute of Navigation*, pp. 730–743.

Dbouk, H. and Schön, S. (2020). Reliable bounding zones and inconsistency measures for GPS positioning using geometrical constraints. *Acta Cybernetica*, 24(3):573–591.

DeCleene, B. (2000). Defining pseudorange integrity-overbounding. In: *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, pp. 1916–1924.

Dosi, G. and Egidi, M. (1991). Substantive and procedural uncertainty: an exploration of economic behaviours in changing environments. *Journal of evolutionary economics*, 1:145–168.

Drevelle, V. and Bonnifait, P. (2009). High integrity GNSS location zone characterization using interval analysis. In: *Proceedings of the 22nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2009)*, pp. 2178–2187.

Drevelle, V. and Bonnifait, P. (2011). A set-membership approach for high integrity height-aided satellite positioning. *GPS Solutions*, 15(4):357–368.

Eldar, Y. C., Beck, A., and Teboulle, M. (2008). A minimax Chebyshev estimator for bounded error estimation. *IEEE transactions on signal processing*, 56(4):1388–1397.

Eppstein, D. (1995). Zonohedra and zonotopes. Technical report, Dept. of Information & Computer Science U.C. Irvine, CA, USA.

European Commission (2016). European GNSS (Galileo) Open Service – Ionospheric Correction Algorithm for Galileo Single Frequency Users. Technical document.

EUSPA (2021). Report on road user needs and requirements: Outcome of the EUSPA user consultation platform, Issue/Rev: 3.0. Technical report, European Union Agency for the Space Programme (EUSPA).

Faes, M. G., Daub, M., Marelli, S., Patelli, E., and Beer, M. (2021). Engineering analysis with probability boxes: A review on computational methods. *Structural Safety*, 93:102092.

Feng, P., Li, F., Yan, J., Zhang, F., and Barriot, J.-P. (2020). Assessment of the accuracy of the Saastamoinen model and VMF1/VMF3 mapping functions with respect to ray-tracing from radiosonde data in the framework of GNSS meteorology. *Remote Sensing*, 12(20):3337.

Ferrero, A. and Salicone, S. (2006). Measurement uncertainty. *IEEE instrumentation & measurement magazine*, 9(3):44–51.

Gallon, E. (2023). *High-Integrity Modeling of Non-Stationary Kalman Filter Input Error Processes and Application to Aircraft Navigation*. Diss., Illinois Institute of Technology.

Gallon, E., Joerger, M., and Pervan, B. (2021a). Development of stochastic IMU error models for INS/GNSS integration. In: *Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021), St. Louis, MI, USA*, pp. 20–24.

Gallon, E., Joerger, M., and Pervan, B. (2021b). Robust modeling of GNSS tropospheric delay dynamics. *IEEE Transactions on Aerospace and Electronic Systems*, 57(5):2992–3003.

Gauss, C. F. (1809). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Sumtibus F. Perthes et I.H. Besser.

Ge, Y., Wang, Z., and Zhu, Y. (2017). Reduced ARAIM monitoring subset method based on satellites in different orbital planes. *GPS Solutions*, 21:1443–1456.

Geweke, J. (1986). Exact inference in the inequality constrained normal linear regression model. *Journal of Applied econometrics*, 1(2):127–141.

Groves, P. D., Jiang, Z., Rudi, M., and Strode, P. (2013). A portfolio approach to NLOS and multipath mitigation in dense urban areas. In: *Proceedings of the 26th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2013)*, pp. 3231–3247.

Hannah, B. M. (2001). *Modelling and simulation of GPS multipath propagation*. Diss., Queensland University of Technology.

Harris, R. B. and Lightsey, E. G. (2009). A general model of multipath error for coherently tracked BOC modulated signals. *IEEE Journal of Selected Topics in Signal Processing*, 3(4):682–694.

Hartinger, H. and Brunner, F. (1999). Variances of GPS phase observations: the SIGMA-ε model. *GPS Solutions*, 2(4):35–43.

Hauschild, A. (2017). *Springer Handbook of Global Navigation Satellite Systems*, Chap. Basic observation equations, pp. 561–582. In: Teunissen and Montenbruck (2017).

Hauschild, A., Montenbruck, O., Thölert, S., Erker, S., Meurer, M., and Ashjaee, J. (2012). A multi-technique approach for characterizing the SVN49 signal anomaly, part 1: receiver tracking and IQ constellation. *GPS Solutions*, 16(1):19–28.

Hazewinkel, M. (Ed.) (1995). *Encyclopaedia of Mathematics*, Volume 1. Springer.

Henk, M., Richter-Gerbert, J., and Ziegler, G. (1997). Basic properties of convex polytopes. In: *Handbook of Discrete and Computational Geometry*, pp. 243–270. CRC Press LLC, Boca Raton.

Hobiger, T. and Jakowski, N. (2017). *Springer Handbook of Global Navigation Satellite Systems*, Chap. Atmospheric signal propagation, pp. 165–193. In: Teunissen and Montenbruck (2017).

Hofmeister, A. and Böhm, J. (2017). Application of ray-traced tropospheric slant delays to geodetic VLBI analysis. *Journal of Geodesy*, 91:945–964.

Hsu, L.-T., Jan, S.-S., Groves, P. D., and Kubo, N. (2015). Multipath mitigation and NLOS detection using vector tracking in urban environments. *GPS Solutions*, 19:249–262.

Huber, P. J. (1981). *Robust statistics*, Volume 523. John Wiley & Sons.

Hüllermeier, E. and Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. *Machine Learning*, 110:457–506.

ICAO (2023). Annex 10 to the Convention on International Civil Aviation - Aeronautical Telecommunications - Volume I - Radio Navigational Aids. International Standards and Recommended Practices, International Civil Aviation Organization. 8th Edition.

Icking, L., Kersten, T., and Schön, S. (2020). Evaluating the urban trench model for improved GNSS positioning in urban areas. In: *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pp. 631–638.

Icking, L., Ruwisch, F., and Schön, S. (2022). Multipath characterization using ray-tracing in urban trenches. In: Freymueller, J. (Ed.), *Geodesy for a Sustainable Earth: Proceedings of the 2021 Scientific Assembly of the International Association of Geodesy, Beijing, China, June 28–July 2, 2021*, pp. 359–365. Springer.

Ingimundarson, A., Bravo, J. M., Puig, V., Alamo, T., and Guerra, P. (2009). Robust fault detection using zonotope-based set-membership consistency test. *International journal of adaptive control and signal processing*, 23(4):311–330.

IS-GPS-200N (2022). NAVSTAR GPS Space Segment/Navigation User Segment Interfaces IS-GPS-200. Interface specification document, MilComm & PNT Directorate, Space Systems Command (SSC).

ISO/IEC GUIDE 98-3:2008(E) (2008). Uncertainty of measurement - Part 3: Guide to the expression of uncertainty in measurement (GUM:1995). Normative document, International Organization for Standardization, Geneva, Switzerland.

Jan, S.-S., Chan, W., Walter, T., and Enge, P. (2001). 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)*, pp. 2366–2375.

Jaulin, L., Kieffer, M., Didrit, O., and Walter, E. (2001). *Applied interval analysis: With examples in parameter and state estimation, robust control and robotics*. Springer, London.

JCGM 100:2008 (2008). Evaluation of measurement data - Guide to the expression of uncertainty in measurement: JCGM 100:2008 (GUM 1995 with minor corrections). Normative document, Joint Committee for Guides in Metrology (JCGM).

JCGM 102:2011 (2011). Evaluation of measurement data - Supplement 2 to the "Guide to the expression of uncertainty in measurement" - extension to any number of output quantities. Normative document, Joint Committee for Guides in Metrology (JCGM).

JCGM 200:2012(E/F) (2012). International vocabulary of metrology - basic and general concepts and associated terms (VIM) 2008 version with minor corrections. Normative document, Joint Committee for Guides in Metrology (JCGM).

JCGM GUM-6:2020 (2020). Guide to the expression of uncertainty in measurement - Part 6: Developing and using measurement models. Normative document, Joint Committee for Guides in Metrology (JCGM).

Jiang, Y. and Wang, J. (2014). A new approach to calculate the vertical protection level in A-RAIM. *The Journal of Navigation*, 67(4):711–725.

Joerger, M., Chan, F.-C., Langel, S., and Pervan, B. (2012). 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)*, pp. 2785–2807.

Joerger, M., Chan, F.-C., and Pervan, B. (2014). Solution separation versus residual-based RAIM. *NAVIGATION: Journal of the Institute of Navigation*, 61(4):273–291.

Joerger, M. and Pervan, B. (2014). Solution Separation and Chi-Squared ARAIM for fault detection and exclusion. In: *2014 IEEE/ION Position, Location and Navigation Symposium (PLANS) 2014*, pp. 294–307.

Joerger, M. and Pervan, B. (2016). Fault detection and exclusion using solution separation and Chi-squared ARAIM. *IEEE Transactions on Aerospace and electronic systems*, 52(2):726–742.

Joerger, M., Stevanovic, S., Chan, F.-C., Langel, S., and Pervan, B. (2013). Integrity risk and continuity risk for fault detection and exclusion using solution separation ARAIM. In: *Proceedings of the 26th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2013)*, pp. 2702–2722.

Johnston, G., Riddell, A., and Hausler, G. (2017). *Springer Handbook of Global Navigation Satellite Systems*, Chap. The international GNSS service, pp. 967–982. In: Teunissen and Montenbruck (2017).

Kaarls, R. (1981). Recommendation INC-1. *BIPM Proc.-Verb. Com. Int. Poids et Mesures*, 49:A1–A12.

Kabir, S., Wagner, C., and Ellerby, Z. (2024). Toward handling uncertainty-at-source in AI – A review and next steps for interval regression. *IEEE Transactions on Artificial Intelligence*, 5(1):3–22.

Kacewicz, B., Milanese, M., Tempo, R., and Vicino, A. (1986). Optimality of central and projection algorithms for bounded uncertainty. *Systems & control letters*, 8(2):161–171.

Kaplan, E. D. and Hegarty, C. (Ed.) (2017). *Understanding GPS/GNSS: Principles and applications*. Artech house.

Karimidoona, A. and Schön, S. (2023). Predicting C/N0 as a key parameter for network RTK integrity prediction in urban environments. *Remote Sensing*, 15(19):4850.

Kaspar, F., Müller-Westermeier, G., Penda, E., Mächel, H., Zimmermann, K., Kaiser-Weiss, A., and Deutschländer, T. (2013). Monitoring of climate change in Germany–data, products and services of Germany's National Climate Data Centre. *Advances in Science and Research*, 10(1):99–106.

Katipamula, S. and Brambley, M. R. (2005). Methods for fault detection, diagnostics, and prognostics for building systems—a review, part i. *Hvac&R Research*, 11(1):3–25.

Kerekes, G. and Schwieger, V. (2020). Elementary error model applied to terrestrial laser scanning measurements: Study case arch dam Kops. *Mathematics*, 8(4):593.

Kesavan, P. and Lee, J. H. (2001). A set based approach to detection and isolation of faults in multivariable systems. *Computers & Chemical Engineering*, 25(7-8):925–940.

Khanafseh, S., Kujur, B., Joerger, M., Walter, T., Pullen, S., Blanch, J., Doherty, K., Norman, L., de Groot, L., and Pervan, B. (2018). GNSS multipath error modeling for automotive applications. In: *Proceedings of the 31st International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2018)*, pp. 1573–1589.

Klobuchar, J. (1987). Ionospheric time-delay algorithm for single-frequency GPS users. *IEEE Transactions on Aerospace and Electronic Systems*, AES-23(3):325–331.

Koch, K.-R. (1999). *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer Science & Business Media.

Kong, S.-H. (2011). Statistical analysis of urban GPS multipaths and pseudo-range measurement errors. *IEEE transactions on aerospace and electronic systems*, 47(2):1101–1113.

Kreinovich, V. (1995). Data processing beyond traditional statistics: applications of interval computations. a brief introduction. *Extended Abstracts of APIC'95 (Supplement to the International Journal of Reliable Computing)*, pp. 13–21.

Kurzhanskiy, A. A. and Varaiya, P. (2007). Ellipsoidal techniques for reachability analysis of discrete-time linear systems. *IEEE Transactions on Automatic Control*, 52(1):26–38.

Kutterer, H. (1994). *Intervallmathematische Behandlung endlicher Unschärfen linearer Ausgleichungsmodelle*. Dissertation, Universität Fridericiana zu Karlsruhe (TH).

Kutterer, H. (1999). On the sensitivity of the results of least-squares adjustments concerning the stochastic model. *Journal of Geodesy*, 73(7):350–361.

Kutterer, H. (2001). Uncertainty assessment in geodetic data analysis. In: *First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS. Swiss Federal Institute of Technology Zurich, Institute of Geodesy and Photogrammetry-Report No. 295*, pp. 7–12, Zürich, Schweiz.

Kutterer, H. (2002). Joint treatment of random variability and imprecision in GPS data analysis. *Journal of Global Positioning Systems*, 1(2):96–105.

Lagler, K., Schindelegger, M., Böhm, J., Krásná, H., and Nilsson, T. (2013). GPT2: Empirical slant delay model for radio space geodetic techniques. *Geophysical research letters*, 40(6):1069–1073.

Lai, Y.-F., Blanch, J., and Walter, T. (2023). Troposphere delay model error analysis with application to vertical protection level calculation. In: *Proceedings of the 2023 International Technical Meeting of The Institute of Navigation*, pp. 903–921.

Landskron, D. and Böhm, J. (2018). VMF3/GPT3: refined discrete and empirical troposphere mapping functions. *Journal of Geodesy*, 92(4):349–360.

Langley, R. B. et al. (1999). Dilution of precision. *GPS world*, 10(5):52–59.

Larson, J. D., Gebre-Egziabher, D., and Rife, J. H. (2019). Gaussian-Pareto overbounding of DGNSS pseudoranges from CORS. *Navigation*, 66(1):139–150.

Le, V. T. H., Stoica, C., Alamo, T., Camacho, E. F., and Dumur, D. (2013). Zonotopic guaranteed state estimation for uncertain systems. *Automatica*, 49(11):3418–3424.

Lee, Y. C. (1986). 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)*, pp.

1–4.

Liew, C. K. (1976). Inequality constrained least-squares estimation. *Journal of the American Statistical Association*, 71(355):746–751.

Liu, B., Gao, Y., Gao, Y., and Wang, S. (2022). HPL calculation improvement for Chi-squared residual-based ARAIM. *GPS Solutions*, 26(2):45.

Luo, X., Mayer, M., and Heck, B. (2009). Improving the stochastic model of GNSS observations by means of SNR-based weighting. In: *Observing our changing earth*, pp. 725–734. Springer.

Macabiau, C., Milner, C., Chabory, A., Suard, N., Rodriguez, C., Mabilleau, M., Vuillaume, J., and Hegron, S. (2015). Nominal bias analysis for ARAIM user. In: *Proceedings of the 2015 International Technical Meeting of The Institute of Navigation*, pp. 713–732.

Macabiau, C., Milner, C., Tessier, Q., Mabilleau, M., Vuillaume, J., Suard, N., and Rodriguez, C. (2014). Impact of nominal bias bounding techniques on final ARAIM user performance. In: *Proceedings of the 2014 International Technical Meeting of The Institute of Navigation*, pp. 68–77.

McGraw, G. A., Young, R. S., Reichenauer, K., Stevens, J., and Ventrone, F. (2004). GPS multipath mitigation assessment of digital beam forming antenna technology in a JPALS dual frequency smoothing architecture. In: *Proceedings of the 2004 National Technical Meeting of the Institute of Navigation*, pp. 561–572.

Mead, J. L. and Renaut, R. A. (2010). Least squares problems with inequality constraints as quadratic constraints. *Linear Algebra and its Applications*, 432(8):1936–1949.

Meizel, D., Lévêque, O., Jaulin, L., and Walter, E. (2002). Initial localization by set inversion. *IEEE transactions on robotics and Automation*, 18(6):966–971.

Milanese, M. (1995). Properties of least-squares estimates in set membership identification. *Automatica*, 31(2):327–332.

Milanese, M. and Vicino, A. (1991). Optimal estimation theory for dynamic systems with set membership uncertainty: An overview. *Automatica*, 27(6):997–1009.

Milner, C. D. and Ochieng, W. Y. (2011). Weighted RAIM for APV: the ideal protection level. *The Journal of Navigation*, 64(1):61–73.

Möller, B. and Beer, M. (2008). Engineering computation under uncertainty–capabilities of non-traditional models. *Computers & Structures*, 86(10):1024–1041.

Moore, R. E., Kearfott, R. B., and Cloud, M. J. (2009). *Introduction to interval analysis*. SIAM.

Morton, Y. J., van Diggelen, F., Spilker Jr, J. J., Parkinson, B. W., Lo, S., and Gao, G. (Ed.) (2021). *Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications*. John Wiley & Sons.

Narayanan, S. (2023). Residual error model to bound unmodeled tropospheric delays for terrestrial navigation systems for very low elevation angles. In: *Proceedings of the 36th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2023)*, pp. 3908–3922.

Nava, B., Coisson, P., and Radicella, S. (2008). A new version of the NeQuick ionosphere electron density model. *Journal of Atmospheric and Solar-Terrestrial Physics*, 70(15):1856–1862.

Neumann, I., Kutterer, H., and Schön, S. (2006). Outlier detection in geodetic applications with respect to observation imprecision. In: *Proceedings of NSF Workshop on Reliable Engineering Computing, Savannah, Georgia*, pp. 75–90.

Nikiforov, I. (2019). From pseudorange overbounding to integrity risk overbounding. *Navigation*, 66(2):417–439.

Ober, P. (1998). RAIM performance: How algorithms differ. In: *Proceedings of the 11th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1998)*, pp. 2021–2030.

Oberkampf, W. L., Helton, J. C., Joslyn, C. A., Wojtkiewicz, S. F., and Ferson, S. (2004). Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering & System Safety*, 85(1-3):11–19.

O'Leary, D. P. and Rust, B. W. (1986). Confidence intervals for inequality-constrained least squares problems, with applications to ill-posed problems. *SIAM journal on scientific and statistical computing*, 7(2):473–489.

Pagot, J.-B., Julien, O., and Gregoire, Y. (2017). Worst impact of pseudorange nominal bias on the position in a civil aviation context. In: *ITSNT 2017, 4th International Technical Symposium on Navigation and Timing*.

Pany, T., Kaniuth, R., and Eissfeller, B. (2005). Deep integration of navigation solution and signal processing. In: *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2005)*, pp. 1095–1102.

Parkinson, B. W. and Axelrad, P. (1988). Autonomous GPS integrity monitoring using the pseudorange residual. *Navigation*, 35(2):255–274.

Parkinson, B. W., Morton, Y. J., van Diggelen, F., and Spilker Jr, J. J. (2021). *Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications*, Chap. Introduction, early history, and assuring PNT (PTA), pp. 1–42. Volume 1 of Morton et al. (2021).

Pelzer, H. (1985). Grundlagen der mathematischen Statistik und der Ausgleichungsrechnung. *Geodätische Netze in Landes-und Ingenieurvermessung*.

Peng, J., Zhang, H., Shong, S., and Guo, C. (2006). An aggregate constraint method for inequality-constrained least squares problems. *Journal of Geodesy*, 79:705–713.

Perea, S., Meurer, M., Rippl, M., Belabbas, B., and Joerger, M. (2017). URA/SISA analysis for GPS and Galileo to support ARAIM. *NAVIGATION: Journal of the Institute of Navigation*, 64(2):237–254.

Pérez, R. N., Amaro, J., and Arriola, E. R. (2016). The low-energy structure of the nucleon–nucleon interaction: statistical versus systematic uncertainties. *Journal of Physics G: Nuclear and Particle Physics*, 43(11):114001.

Phelts, R. E., Blanch, J., Walter, T., and Enge, P. (2014). The effect of nominal signal deformation biases on ARAIM users. In: *Proceedings of the 2014 International Technical Meeting of The Institute of Navigation*, pp. 56–67.

Pons, R., Jauberthie, C., Travé-Massuyès, L., and Goupil, P. (2008). Interval analysis based learning for fault model identification: Application to control surfaces oscillatory failures. In: *Proceedings of the 22nd International Workshop on Qualitative Reasoning*, pp. 115–122.

Potter, J. and Suman, M. (1977). Thresholdless redundancy management with arrays of skewed instruments. *Integrity in Electronic Flight Control Systems*, (224):11–15.

Puig, V. (2010). Fault diagnosis and fault tolerant control using set-membership approaches: Application to real case studies. *International Journal of Applied Mathematics and Computer Science*, 20(4):619–635.

Puig, V. and Blesa, J. (2013). Limnimeter and rain gauge FDI in sewer networks using an interval parity equations based detection approach and an enhanced isolation scheme. *Control Engineering Practice*, 21(2):146–170.

Puig, V., Witczak, M., Nejjari, F., Quevedo, J., and Korbicz, J. (2007). A GMDH neural network-based approach to passive robust fault detection using a constraint satisfaction backward test. *Engineering Applications of Artificial Intelligence*, 20(7):886–897.

Pullen, S. and Joerger, M. (2021). *Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications*, Chap. GNSS Integrity and Receiver Autonomous Integrity Monitoring (RAIM), pp. 591–617. Volume 1 of Morton et al. (2021).

Racelis, D. and Joerger, M. (2022). Evaluating new and current horizontal protection levels in the baseline ARAIM algorithm. In: *Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022)*, pp. 238–246.

Reid, T. G., Houts, S. E., Cammarata, R., Mills, G., Agarwal, S., Vora, A., and Pandey, G. (2019). Localization requirements for autonomous vehicles. *arXiv preprint arXiv:1906.01061*.

Reid, T. G., Neish, A., et al. (2023). Localization & mapping requirements for level 2+ autonomous vehicles. In: *Proceedings of the 2023 International Technical Meeting of The Institute of Navigation*, pp. 107–123.

Rife, J., Pullen, S., Enge, P., and Pervan, B. (2006). Paired overbounding for nonideal LAAS and WAAS error distributions. *IEEE Transactions on Aerospace and Electronic Systems*, 42(4):1386–1395.

Rife, J., Pullen, S., Pervan, B., and Enge, P. (2004a). Paired overbounding and application to GPS augmentation. In: *2004 IEEE position, location and navigation symposium*, pp. 439–446. IEEE.

Rife, J., Walter, T., and Blanch, J. (2004b). Overbounding SBAS and GBAS error distributions with excess-mass functions. In: *Proceedings of the International Symposium on GNSS/GPS, Sydney, Australia*, pp. 6–8.

Ritter, J. (1990). An efficient bounding sphere. *Graphics gems*, 1:301–303.

Roese-Koerner, L., Devaraju, B., Sneeuw, N., and Schuh, W.-D. (2012). A stochastic framework for inequality constrained estimation. *Journal of Geodesy*, 86:1005–1018.

Rohou, S., Jaulin, L., Mihaylova, L., Le Bars, F., and Veres, S. M. (2020). *Reliable robot localization: a constraint-programming approach over dynamical systems*. John Wiley & Sons.

Rózsa, S. (2018). A new approach for assessing tropospheric delay model performance for safety-of-life GNSS applications. In: Heck, A., Seitz, K., Grombein, T., Mayer, M., Stövhase, J.-M., Sumaya, H., Wampach, M., Westerhaus, M., Dalheimer, L., and Senger, P. (Ed.), *(Schw)Ehre, wem (Schw)Ehre gebührt : Festschrift zur Verabschiedung von Prof. Dr.-Ing. Dr. h.c. Bernhard Heck*. KIT Scientific Publishing.

Rózsa, S., Ambrus, B., Juni, I., Ober, P. B., and Mile, M. (2020). An advanced residual error model for tropospheric delay estimation. *GPS Solutions*, 24(4):1–15.

RTCA/DO-229D (2006). Minimum operational performance standards for Global Positioning System/Wide Area Augmentation System airborne equipment. Standard specification, Radio Technical Commission for Aeronautics (RTCA) Special Committee 159 (SC-159), Washington, D.C., USA.

RTCA/DO-229F (2020). Minimum operational performance standards (MOPS) for Global Positioning System/Satellite-Based Augmentation System airborne equipment. Standard specification, Radio Technical Commission for Aeronautics (RTCA) Special Committee 159 (SC-159), Washington, D.C., USA.

Ruwisch, F. and Schön, S. (2022). Performance assessment of GNSS RTK positioning in urban environments: Outlier detection versus 3DMA-FDE. In: *Proceedings of the 35th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2022)*, pp. 2649–2663.

Ruwisch, F. and Schön, S. (2024). Feature map aided robust high precision GNSS positioning in harsh urban environments. Conference presentation. In: *The 35th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2022)*.

Saastamoinen, J. (1972). Atmospheric correction for the troposphere and stratosphere in radio ranging satellites. *The use of artificial satellites for geodesy*, 15:247–251.

Samada, S. E., Puig, V., and Nejjari, F. (2023). Zonotopic recursive least-squares parameter estimation: Application to fault detection. *International Journal of Adaptive Control and Signal Processing*, 37(4):993–1014.

Sánchez, J. S., Gerhmann, A., Thevenon, P., Brocard, P., Afia, A. B., and Julien, O. (2016). Use of a FishEye camera for GNSS NLOS exclusion and characterization in urban environments. In: *Proceedings of the 2016 International Technical Meeting of The Institute of Navigation*, pp. 283–292.

Schempp, T. R. and Rubin, A. L. (2002). An application of Gaussian overbounding for the WAAS fault free error analysis. In: *Proceedings of the 15th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2002)*, pp. 766–772.

Schön, S. (2003). *Analyse und Optimierung geodätischer Messanordnungen unter besonderer Berücksichtigung des Intervallansatzes*. Dissertation, Universität Fridericiana zu Karlsruhe (TH).

Schön, S., Brenner, C., Alkhatib, H., Coenen, M., Dbouk, H., Garcia-Fernandez, N., Fischer, C., Heipke, C., Lohmann, K., Neumann, I., Nguyen, U., Paffenholz, J.-A., Peters, T., Rottensteiner, F., Schachtschneider, J., Sester, M., Sun, L., Vogel, S., Voges, R., and Wagner, B. (2018). Integrity and collaboration in dynamic sensor networks. *Sensors*, 18(7).

Schön, S. and Kutterer, H. (2001a). Interval-based description of measurement uncertainties and network optimization. In: *Proceedings of the First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS. Swiss Federal Institute of Technology Zurich, Institute of Geodesy and Photogrammetry-Report No. 295*, pp. 41–46, Zürich, Schweiz.

Schön, S. and Kutterer, H. (2001b). Optimal design of geodetic monitoring networks by means of interval mathematics. In: *Proceedings of the 10th FIG International Symposium on Deformation Analysis*, pp. 41–46.

Schön, S. and Kutterer, H. (2005a). Realistic uncertainty measures for GPS observations. In: Sansò, F. (Ed.), *A Window on the Future of Geodesy*, Volume 128 of *International Association of Geodesy Symposia*, pp. 54–59. Springer-Verlag, Berlin/Heidelberg.

Schön, S. and Kutterer, H. (2005b). Using zonotopes for overestimation-free interval least-squares–some geodetic applications. *Reliable Computing*, 11(2):137–155.

Schön, S. and Kutterer, H. (2006). Uncertainty in GPS networks due to remaining systematic errors: The interval approach. *Journal of Geodesy*, 80(3):150–162.

Schweppe, F. (1968). Recursive state estimation: Unknown but bounded errors and system inputs. *IEEE Transactions on Automatic Control*, 13(1):22–28.

Schwieger, V. (1999). *Ein Elementarfehlermodell für GPS-Überwachungsmessungen: Konstruktion und Bedeutung interepochaler Korrelationen.* Schriftenreihe der Fachrichtung Vermessungswesen der Universität Hannover.

Schwieger, V. (2007a). Determination of synthetic covariance matrices—An application to GPS monitoring measurements. In: *2007 15th European Signal Processing Conference*, pp. 1161–1165. IEEE.

Schwieger, V. (2007b). Sensitivity analysis as a general tool for model optimisation – examples for trajectory estimation. *Journal of Applied Geodesy*, 1(1):27–34.

Scott, J. K., Raimondo, D. M., Marseglia, G. R., and Braatz, R. D. (2016). Constrained zonotopes: A new tool for set-based estimation and fault detection. *Automatica*, 69:126–136.

Seidel, R. (1991). Small-dimensional linear programming and convex hulls made easy. *Discrete & Computational Geometry*, 6:423–434.

She, J., Misovec, K., Blanch, J., Caccioppoli, N., Duchet, D., Tijero, E. D., Liu, F., Racelis, D., Joerger, M., and Sgammini, M. (2023). Implementation of the reference Advanced RAIM user algorithm. In: *Proceedings of the 36th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2023)*, pp. 1099–1127.

Shi, X., Mao, G., Anderson, B. D., Yang, Z., and Chen, J. (2017). Robust localization using range measurements with unknown and bounded errors. *IEEE Transactions on Wireless Communications*, 16(6):4065–4078.

Skyum, S. (1991). A simple algorithm for computing the smallest enclosing circle. *Information Processing Letters*, 37(3):121–125.

Smyrnaios, M., Schön, S., and Nicolás, M. L. (2013). Multipath propagation, characterization and modeling in gnss. In: Jin, S. (Ed.), *Geodetic Sciences - Observations, Modeling and Applications*, Chap. 2, pp. 99–125. IntechOpen, Rijeka.

Soatti, G., Nicoli, M., Garcia, N., Denis, B., Raulefs, R., and Wymeersch, H. (2018). Implicit cooperative positioning in vehicular networks. *IEEE Transactions on Intelligent Transportation Systems*, 19(12):3964–3980.

Su, J. and Schön, S. (2021). Improved observation interval bounding for multi-GNSS integrity monitoring in urban navigation. In: *Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*, pp. 4141–4156.

Su, J. and Schön, S. (2022a). Advances in deterministic approaches for bounding uncertainty and integrity monitoring of autonomous navigation. In: *Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022)*, pp. 1442–1454.

Su, J. and Schön, S. (2022b). Bounding the residual tropospheric error by interval analysis. In: *International Association of Geodesy Symposia*, pp. 1–10. Springer Berlin Heidelberg.

Su, J. and Schön, S. (2022c). Deterministic approaches for bounding GNSS uncertainty: A comparative analysis. In: *2022 10th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC)*, pp. 1–8.

Su, J., Schön, S., and Gallon, E. (2024). Reliable overbounding for stochastic IMU error models using interval analysis. In: *Proceedings of the 37th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2024)*, pp. 1828–1842.

Su, J., Schön, S., and Joerger, M. (2023). Towards a set-based detector for GNSS integrity monitoring. In: *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pp. 421–429.

Tannert, C., Elvers, H.-D., and Jandrig, B. (2007). The ethics of uncertainty: In the light of possible dangers, research becomes a moral duty. *EMBO reports*, 8(10):892–896.

Teunissen, P., Ciuban, S., Yin, C., Noort, B. v., Zaminpardaz, S., and Tiberius, C. (2024). The DIA-estimator for positional integrity: Design and computational challenges. In: *International Association of Geodesy Symposia*, pp. 1–8. Springer Berlin Heidelberg.

Teunissen, P. J. (1990). An integrity and quality control procedure for use in multi sensor integration. In: *Proceedings of the 3rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1990)*, pp. 513–522.

Teunissen, P. J. (2003). *Adjustment theory.* VSSD Netherlands, Delft, The Netherlands.

Teunissen, P. J. (2006). *Testing theory: an introduction.* VSSD Netherlands, Delft, The Netherlands.

Teunissen, P. J. (2018). Distributional theory for the DIA method. *Journal of Geodesy*, 92(1):59–80.

Teunissen, P. J. and Montenbruck, O. (2017). *Springer Handbook of Global Navigation Satellite Systems.* Springer Cham.

Traub, J. F., , Wasikowski, G., and Wozinakowski, H. (1988). *Information-based complexity.* Kluwer Academic Publishers, New York, USA.

U.S. Standard Atmosphere, 1976 (1976). U.S. Standard Atmosphere, 1976. Standard, National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration, United States Air Force, Washington, D.C., USA.

Van Nee, R. D. (1993). Spread-spectrum code and carrier synchronization errors caused by multipath and interference. *IEEE transactions on aerospace and electronic systems*, 29(4):1359–1365.

VMF Data Server (2020). Online Ray-Tracer. https://vmf.geo.tuwien.ac.at/raytracer.html. Last updated: 12/07/2020. Accessed: 30/08/2021.

Walter, T. (2017). *Springer Handbook of Global Navigation Satellite Systems*, Chap. Satellite based augmentation systems, pp. 339–361. In: Teunissen and Montenbruck (2017).

Walter, T. and Blanch, J. (2015). Characterization of GNSS clock and ephemeris errors to support ARAIM. In: *Proceedings of the ION 2015 Pacific PNT Meeting*, pp. 920–931.

Walter, T., Blanch, J., and Enge, P. (2014). Reduced subset analysis for multi-constellation ARAIM. In: *Proceedings of the 2014 International Technical Meeting of The Institute of Navigation*, pp. 89–98.

Walter, T. and Enge, P. (1995). Weighted RAIM for precision approach. In: *Proceedings of ION GPS*, Volume 8, pp. 1995–2004.

Wang, J. and Ober, P. B. (2009). On the availability of fault detection and exclusion in GNSS receiver autonomous integrity monitoring. *the Journal of Navigation*, 62(2):251–261.

Wang, N., Li, Z., Huo, X., Li, M., Yuan, Y., and Yuan, C. (2019). Refinement of global ionospheric coefficients for GNSS applications: Methodology and results. *Advances in Space Research*, 63(1):343–358.

Wang, N., Yuan, Y., Li, Z., and Huo, X. (2016). Improvement of Klobuchar model for GNSS single-frequency ionospheric delay corrections. *Advances in Space Research*, 57(7):1555–1569.

Wang, R. and Walter, T. (2023). Characterization and comparison of Galileo and GPS anomalies. In: *Proceedings of the 2023 International Technical Meeting of The Institute of Navigation*, pp. 597–610.

Wang, S., Zhai, Y., Chi, C., Zhan, X., and Jiang, Y. (2023). Implementation and analysis of fault grouping for multi-constellation advanced RAIM. *Advances in Space Research*, 71(11):4765–4786.

Wang, Y., Chen, X., and Liu, P. (2018). Statistical multipath model based on experimental GNSS data in static urban canyon environment. *Sensors*, 18(4):1149.

Watkins, J. and Yurkovich, S. (1996). Fault detection using set-membership identification. *IFAC Proceedings Volumes*, 29(1):3993–3998.

Welzl, E. (1991). Smallest enclosing disks (balls and ellipsoids). In: *New Results and New Trends in Computer Science: Graz, Austria, June 20–21, 1991 Proceedings*, pp. 359–370. Springer.

Wen, W. W., Zhang, G., and Hsu, L.-T. (2019). GNSS NLOS exclusion based on dynamic object detection using LiDAR point cloud. *IEEE transactions on intelligent transportation systems*, 22(2):853–862.

Wendel, J. (2022). GNSS pseudorange fault detection and exclusion with multiple outliers. In: *Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022)*, pp. 1481–1495.

Werner, H. J. (1990). On inequality constrained generalized least-squares estimation. *Linear Algebra and its Applications*, 127:379–392.

WG-C ARAIM TSG (2012). ARAIM technical subgroup milestone 1 report. Technical report, EU-US Cooperation on Satellite Navigation.

WG-C ARAIM TSG (2015). ARAIM technical subgroup milestone 2 report. Technical report, EU-US Cooperation on Satellite Navigation.

WG-C ARAIM TSG (2016). ARAIM technical subgroup milestone 3 report. Technical report, EU-US Cooperation on Satellite Navigation.

WG-C ARAIM TSG (2019). Advanced RAIM reference airborne algorithm description document v3.1. Technical document.

WG-C ARAIM TSG (2022). Advanced RAIM reference airborne algorithm description document v4.2 draft. Technical document.

Xu, L. and Rife, J. (2019). NLOS and multipath detection using Doppler shift measurements. In: *Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019)*, pp. 4064–4075.

Yang, T.-Y. and Sun, D. (2020). Global Navigation Satellite Systems fault detection and exclusion: A parameterized quadratic programming approach. *IEEE Transactions on Aerospace and Electronic Systems*, 56(4):2862–2871.

Young, R. S., McGraw, G. A., and Driscoll, B. T. (1996). Investigation and comparison of horizontal protection level and horizontal uncertainty level in FDE algorithms. In: *Proceedings of the 9th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1996)*, pp. 1607–1614.

Yuan, L., Jin, S., and Hoque, M. (2020). Estimation of LEO-GPS receiver differential code bias based on inequality constrained least square and multi-layer mapping function. *GPS Solutions*, 24(2):57.

Zhai, Y. (2018). *Ensuring navigation integrity and continuity using multi-constellation GNSS*. Dissertation, Illinois Institute of Technology.

Zhai, Y., Joerger, M., and Pervan, B. (2015). Continuity and availability in dual-frequency multi-constellation ARAIM. In: *Proceedings of the 28th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015)*, pp. 664–674.

Zhai, Y., Joerger, M., and Pervan, B. (2018). Fault exclusion in multi-constellation Global Navigation Satellite Systems. *The Journal of Navigation*, 71(6):1281–1298.

Zhang, D., Guo, J., Chen, M., Shi, J., and Zhou, L. (2016). Quantitative assessment of meteorological and tropospheric Zenith Hydrostatic Delay models. *Advances in Space Research*, 58(6):1033–1043.

Zhang, G., Wen, W., and Hsu, L.-T. (2018). A novel GNSS based V2V cooperative localization to exclude multipath effect using consistency checks. In: *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pp. 1465–1472.

Zhang, K. and Papadimitratos, P. (2021). Fast multiple fault detection and exclusion (FM-FDE) algorithm for standalone GNSS receivers. *IEEE Open Journal of the Communications Society*, 2:217–234.

Zhu, C., Meurer, M., and Joerger, M. (2023). Integrity analysis for greedy search based fault exclusion with a large number of faults. In: *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pp. 430–435.

Zhu, J.-j., Santerre, R., and Chang, X.-W. (2005). A bayesian method for linear, inequality-constrained adjustment and its application to GPS positioning. *Journal of Geodesy*, 78:528–534.

Ziegler, G. M. (1995). *Lectures on polytopes*, Volume 152 of *Graduate texts in mathematics*. Springer, New York.

Acknowledgement

This thesis was developed during my employment at the Institut für Erdmessung (IfE) at the Leibniz Universität Hannover and supported by the German Research Foundation (DFG) as a part of the Research Training Group (RTG) i.c.sens [GRK2159].

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Prof. Steffen Schön, for his exceptional guidance and unwavering support throughout my entire PhD journey. His profound expertise and insightful perspectives have constantly inspired my curiosity, shaped my research, and motivated my progress. His kindness and encouragement have been instrumental in building my confidence as a researcher. I'm especially grateful for his support in facilitating my participation in multiple international conferences, a summer school, and a research stay abroad in the US. These experiences have broadened my knowledge, provided valuable networking opportunities, and helped me gain insights from worldwide professionals.

I am also deeply thankful to Prof. Mathieu Joerger for hosting me in his lab at Virginia Tech and for the many constructive discussions we shared — in Hannover, Monterey, Blacksburg, Baltimore, and over Zoom. The development of the integrity monitoring algorithms in this thesis would not have been possible without his collaboration and inspiration. I also extend my sincere thanks to Prof. Joerger and Prof. Franz Rottensteiner, my thesis co-reviewers, for their time and thoughtful feedback.

I would like to thank all my colleagues at IfE for their support and collaboration. Special thanks go to Dr. Tobias Kersten for providing the LaTeX template, to Ahmed Elmaghraby, Kai Baasch, Yichen Liu (TU Munich), Rui Wang (University of Stuttgart), and Chengyu Yin (TU Delft) for proofreading and providing valuable feedback on the thesis, and to Qianwen Lin, Dennis Kulemann and Thomas Maschke for their help in preparing and supporting my dissertation defense.

I am thankful to all my friends in Hannover and beyond, whose companionship helped me overcome challenges, manage stress, and share moments of joy along the way.

Finally, and most importantly, I extend my deepest gratitude to my parents for their unwavering love and support throughout this journey.

Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover

(Eine vollständige Liste der Wiss. Arb. ist beim Geodätischen Institut, Nienburger Str. 1, 30167 Hannover erhältlich.)

Nr. 383	KRUSE, Christian:	Impact maps from bomb craters detected in aerial wartime images using marked point processes (Diss. 2023)
Nr. 384	ZOURLIDOU, Stefania:	Traffic Regulation Recognition from GPS Data (Diss. 2023)
Nr. 385	SLEDZ, Artuom:	Thermal anomaly detection based on information fusion from optical and infrared images (Diss. 2023)
Nr. 386	WITTICH, Dennis:	Deep Domain Adaptation for the Pixel-wise Classification of Aerial and Satellite Images (Diss. 2023)
Nr. 387	SINGH, Vishwa Vijay:	Lunar Laser Ranging - Improved Modelling and Parameter Estimation (Diss. 2023)
Nr. 388	HARTMANN, Jens:	Hochgenaue 3D-Erfassung von Großstrukturen durch kinematisches terrestrisches Laserscanning (Diss. 2023)
Nr. 389	ZHUGE, Xia:	Characterizing slope instability kinematics by integrating multi-sensor satellite remote sensing observations (Diss. 2023)
Nr. 390	DOROZYNSKI, Mareike Marianne:	Image Classification and Retrieval in the Context of Silk Heritage using Deep Learning (Diss. 2023)
Nr. 391	KNABE, Annike:	New Concepts for Gravity Field Recovery using Satellites (Diss. 2023)
Nr. 392	KALIA, Andre:	Landslide activity detection based on nationwide Sentinel-1 PSI datasets (Diss. 2023)
Nr. 393	BROCKMEYER, Marco:	Modellierung von Bodenbewegungen anhand heterogener Messverfahren am Beispiel der niedersächsischen Landesfläche (Diss. 2023)
Nr. 394	ZHANG, Mingyue:	Characteristics and Benefits of Differential Lunar Laser Ranging (Diss. 2023)
Nr. 395	DENNIG, Dirk:	Entwicklung eines kinematischen Profilvermessungssystems am Beispiel Kranbahnvermessung (Diss. 2024)
Nr. 396	FUEST, Stefan:	Nudging travelers to societally favorable routes by means of cartographic symbolization (Diss. 2024)
Nr. 397	MOFTIZADEH, Rozhin:	Advanced Particle Filtering for Vehicle Navigation based on Collaborative Information (Diss. 2024)
Nr. 398	VASSILEVA, Magdalena Stefanova:	Satellite Radar Interferometry for Geohazards: from ground deformation to processes understanding (Diss. 2024)
Nr. 399	MALINOVSKAYA, Anna:	Statistical Process Monitoring of Networks (Diss. 2024)
Nr. 400	BANNERT, Jörn:	Der Einfluss von Straßenverkehrslärm und Umgehungsstraßen auf Grundstückswerte in Ortslagen - Bestimmung mittels Expertenbefragung nach der Delphi-Methode (Diss. 2024)
Nr. 401	AXMANN: Jeldrik:	Maximum consensus localization using LiDAR (Diss. 2024)
Nr. 402	TENNSTEDT, Benjamin:	Concept and Evaluation of a Hybridization Scheme for Atom Interferometers and Inertial Measurement Units (Diss. 2024)
Nr. 403	HAKE, Frederic:	Schadenserkennung an Bauwerken mittels maschinellem Lernens (Diss. 2025)
Nr. 404	KARIMIDOONA, Ali:	On Integrity Prediction for Network-RTK Positioning in Urban Environments (Diss. 2025)
Nr. 405	ORTEGA, Mabel:	Domain Adaptation for Deforestation Detection in Remote Sensing: Addressing Class Imbalance and Performance Estimation (Diss. 2025)
Nr. 406	KUPRIYANOV, Alexey:	Investigation of Optical Accelerometry and Novel Satellite Formations for Future Gravimetry Missions (Diss. 2025)
Nr. 407	BREVA, Yannick:	On the Observation Quality of Robot-based GNSS Antenna Calibration for Determining Codephase Corrections (Diss. 2025)
Nr. 408	GUO, Zelong:	Co- and Post-seismic Slip Models Inferred from InSAR Geodesy (Diss. 2025)
Nr. 409:	YUAN, Yunshuang:	Collective Perception: A Fully-Sparse Deep Learning Framework for Multi-Agent Data Fusion (Diss. 2025)
Nr. 410	SU, Jingyao:	Towards interval-based autonomous integrity monitoring: Error bounding and uncertainty propagation (Diss. 2025)