Jackknife Test for Faulty GNSS Measurements Detection under Non-Gaussian Noises

Penggao Yan
Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hong Kong

BIOGRAPHY

Penggao Yan is currently a Ph.D. candidate at the Hong Kong Polytechnic University. He received the bachelor's degree in Communication Engineering in 2018 and the master's degree in Pattern Recognition and Intelligent Systems in 2021, both from Wuhan University, China. His research interests include non-Gaussian noise modeling, fault detection and integrity monitoring in navigation systems, and control-aided localization.

ABSTRACT

Fault detection is crucial to ensure the reliability of navigation systems. However, mainstream fault detection methods are developed based on Gaussian noise assumptions, while other methods targeting non-Gaussian noises lack rigorous statistical properties. The performance and reliability of these methods are challenged in real-world applications. This paper proposes a fault detection method for linearized pseudorange-based positioning systems under non-Gaussian noises. Specifically, this paper proposes a test statistic based on the jackknife technique, which is proved to be the linear combination of measurement noises without any assumption about noise distribution. Furthermore, a hypothesis test with the Bonferroni correction is constructed to detect potential faults in measurements. In a worldwide simulation, the proposed method demonstrates superior performance than the multiple hypothesis solution separation (MHSS) method under non-Gaussian noises. The reliability of the proposed method is further examined in detecting artificially injected faults for a differential global navigation satellite system (DGNSS) positioning system. Moreover, a real-world application to detect satellite clock anomalies for a single point positioning (SPP) system is investigated. The results show a significant improvement in reducing detection delay (8 minutes earlier than MHSS).

I. INTRODUCTION

Gaussian-based fault detection methods are not enough: Fault detection is essential for localization and navigation systems in some safety-critical applications (Joerger & Pervan, 2016; Liu & Morton, 2022; Osechas et al., 2012; B. S. Pervan et al., 1998; R. Wang et al., 2016), which is the technology to check the occurrence of faults in the system as well as to determine the time of fault occurs (Gao et al., 2015). For pseudorange-based positioning systems, such as the single point positioning (SPP) and the differential global navigation satellite system (DGNSS) positioning, the fundamental principle of fault detection is to identify measurements that deviate from the system's expected behavior, where abnormal measurements usually show a considerable inconsistency with normal measurement (Hsu et al., 2017). Based on this principle, fault detection methods such as the range comparison method (Lee, 1986), parity space (B. S. Pervan et al., 1996; Sturza, 1988), chi-squared test (Joerger & Pervan, 2013; Walter & Enge, 1995), and solution separation (Blanch et al., 2010; Brown & McBurney, 1988), are developed and widely applied in satellite-based navigation systems for fault detection and integrity monitoring purposes (Blanch et al., 2015; Joerger et al., 2014). However, a common assumption has been made in developing these methods that the nominal measurement error is Gaussian distributed. Indeed, this assumption brings several benefits and charming properties in constructing and applying these fault detection methods, such as 1) uncorrelated Gaussian variables are independent, which is the foundation for determining the theoretical threshold of the parity space (Sturza, 1988) and the chi-squared methods (Walter & Enge, 1995); 2) the linear combination of Gaussian variables is still Gaussian distributed, which makes it much easy to project measurement domain errors to position domain errors and thus favorite in integrity monitoring (Blanch et al., 2010). However, measurement errors or noises in the real world usually have non-Gaussian properties. For example, the noise of global navigation satellite system (GNSS) signals exhibits a strong non-Gaussian property, as exemplified in many studies (Braff & Shively, 2005; Niu et al., 2014; B. Pervan et al., 2000; Rife, Pullen, & Pervan, 2004; Zhao et al., 2020). The unrealistic Gaussian assumptions can result in increased false alarm rates and degraded fault detection rates in real-world applications, limiting the reliability and effectiveness of preventing systems from faults.

Non-Gaussian fault detection methods are underexplored: A fundamental solution to the non-Gaussian problems is to incorporate non-Gaussian properties into the process of state estimation and fault detection. Although extensive methods have been proposed for state estimation with non-Gaussian noises, such as robust estimation (Pfeifer et al., 2021; Sünderhauf et al., 2013; Wen et al., 2021), particle filters (Gabela et al., 2021; Gupta et al., 2022; Ray et al., 2018), and adaptive error modeling (Pfeifer & Protzel, 2019), little attention is given to fault detection problems with non-Gaussian noises. The early exploration

of non-Gaussian fault detection methods can be roughly classified into two categories: the Gaussian sum filter (GSF) approach and the particle filter (PF) approach. In the first category, Yun et al. (2008) use the Gaussian mixture model (GMM) to model measurement noises and deploy several parallel Kalman filters to deal with each Gaussian component. The fault detection process is realized by comparing the one-side tail probability of the GMM-distributed residual obtained from the GSF with a threshold, which is heuristically obtained by constructing a Kalman filter based detector under Gaussian assumptions. J. Wang et al. (2022) also develop a similar algorithm. The difference is that Wang's approach sums up the residual of each Kalman filter according to the mixture weight and subsequently takes the summation for a Chi-squared test. However, an assumption is made yet not proved in J. Wang et al. (2022) that the weighted summed residual is non-central chi-square distributed. In the second category, a common logic is applied in most of the studies. First, the measurement set is divided into several subsets within which one measurement is excluded (similar to the concept of solution separation). Then, for the full measurement set and each subset, a PF is applied to estimate the state and the likelihood of the predicted state. Finally, a cumulative log-likelihood ratio (LLR) test is constructed by using the likelihood produced by each auxiliary PF and the main PF. Differences among PF-based fault detection literature mainly lie in the determination of the threshold. Rosihan et al. (2006) set the threshold with an empirical value; He et al. (2016) develop a heuristic approach based on the genetic algorithm to solve the near-optimal solution for the threshold; E. Wang et al. (2018) conduct a simulation to determine the threshold. As can be seen, these early explorations of non-Gaussian fault detection methods are either heuristic or lack rigorous statistical properties.

Connections to previous work and contribution in this paper: Previously (Yan, Wen, et al., 2024), we attempted to propose a rigorous fault detection method by establishing a clear relationship between GMM-distributed noises and residuals in the localization system through error propagation, based on which a simple test statistic is constructed for a Chi-squared test. Although this method exemplifies the benefits of incorporating non-Gaussian noise modeling into fault detection methods, it still makes certain assumptions about the distribution of the test statistic (i.e., approaches to a chi-squared distribution), unavoidably compromising the rigor of its statistical properties. In this paper, the relationship between non-Gaussian noises and the test statistic is further investigated. By introducing the jackknife technique, a cross-validation technique in statistics (Quenouille, 1956; Tukey, 1958), this paper formalizes a rigorous hypothesis testing method to detect faulty measurements in localization systems under non-Gaussian noises. The basic idea is to quantify the inconsistency between the observed measurement and the predicted measurement based on subset solutions. Specifically, this paper

1. develops the standard jackknife test (JT) detector with the Bonferroni correction (Bonferroni, 1936) to detect faults in linearized pseudorange-based positioning systems with Gaussian noises (Section II.2);
2. extends the JT detector to non-Gaussian noises by constructing a new test statistic and proving that the constructed test statistic is the linear combination of measurement noises, providing a solid theoretical foundation for fault detection under non-Gaussian noises (Section II.3);
3. in a worldwide simulation, demonstrates the equivalent performance between the JT detector and the multiple hypothesis solution separation (MHSS) method (Blanch et al., 2010; Brown & McBurney, 1988) in detecting faulty measurements for the dual-frequency SPP system under Gaussian noise setting (Section IV.1) and shows the superior performance of the extended JT detector under non-Gaussian noise setting (Section IV.2);
4. applies the proposed method in a DGNSS positioning system to detect artificially injected faults, statistically showing its improved detection rate compared to the MHSS method (Section V);
5. applies the proposed method in a single-frequency SPP system to detect real-world satellite clock anomalies, showing that the proposed method can yield a significantly shorter detection delay and higher detection rate than the MHSS method (Section VI).

This paper limits the scope to linearized pseudorange-based positioning systems, such as linearized SPP and DGNSS positioning systems. In addition, faults in this paper refer to the unmodelled error in the system, such as satellite clock faults in SPP and nonline-of-sight faults in DGNSS.

The contributions of this paper are two folds:

1. Propose the jackknife detector, which provides the theoretical foundation for detecting faults in linearized pseudorange-based positioning systems under non-Gaussian noises;
2. Exemplifies the feasibility and outstanding performance of the jackknife detector through both simulated and real-world experiments.

To the best of the author's knowledge, this is the first work of fault detection for linearized pseudorange-based positioning systems under non-Gaussian noises with rigorous statistical properties.

II. JACKKNIFE TEST FOR LINEAR SYSTEM

1. Linearization of Pseudorange-based Positioning Systems

The GNSS pseudorange measurement model is a non-linear system, which can be formalized as follows (Misra & Enge, 2006):

$$\varrho_i = \sqrt{\left(p_x^i - x\right)^2 + \left(p_y^i - y\right)^2 + \left(p_z^i - z\right)^2} + c\delta_r + \eta_i \,, \tag{1}$$

where ϱ_i is the ith pseudorange measurement, $\mathbf{p}^i = \left[p_x^i, p_y^i, p_z^i\right]^T$ is the position of ith satellite, x, y, z are the receiver position in the Earth-Centered, Earth-Fixed (ECEF) coordinate system, δ_r is the receiver clock bias from single satellite constellation, $c = 3 \times 10^8$ m/s is the speed of light, and η_i is the measurement error.

In most GNSS positioning applications, such as SPP and DGNSS positioning, the pseudorange measurement model is linearized by taking the first-order Taylor expansion at a certain linearized point $\mathbf{x}_0$. A generalized linear system for GNSS positioning can be written as

$$\mathbf{y} = \mathbf{H}\Delta\mathbf{x} + \boldsymbol{\varepsilon} \,, \tag{2}$$

where

$$\mathbf{y} = \begin{bmatrix} f\left(\rho_1, \mathbf{x}_0\right) \\ \vdots \\ f\left(\rho_n, \mathbf{x}_0\right) \end{bmatrix} , \mathbf{H} = \begin{bmatrix} \mathbf{h}\left(\{\mathbf{p}^{1,j}\}, \mathbf{x}_0\right) \\ \vdots \\ \mathbf{h}\left(\{\mathbf{p}^{n,j}\}, \mathbf{x}_0\right) \end{bmatrix} , \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} , \Delta\mathbf{x}{=}\mathbf{x} - \mathbf{x}_0 \,, \tag{3}$$

$f\left(\rho_i, \mathbf{x}_0\right)$ is a function of ith measurement ρ_i (note that ρ_i refers to a generalized measurement, not limited to the pseudorange measurement) and the linearized point $\mathbf{x}_0$; $\mathbf{h}\left(\{\mathbf{p}^{i,j}\}, \mathbf{x}_0\right)$ is a vector function of the collection of satellite positions $\{\mathbf{p}^{i,j}\}$ related to ith measurement and $\mathbf{x}_0$; ε_i is the ith measurement error; and $\mathbf{x}$ is the receiver positioning state (an $m \times 1$ vector). For the SPP system, $\{\mathbf{p}^{i,j}\}$ involves one satellite; for the DGNSS positioning system, $\{\mathbf{p}^{i,j}\}$ involves two satellites. The exact form of Equation (3) for SPP and DGNSS positioning systems is given in Appendix A. In the following, this paper utilizes this general expression to develop the jackknife test for detecting faulty measurements in pseudorange-based positioning systems.

2. Standard Jackknife Test Detector under Gaussian Noises

In statistics, the jackknife is a cross-validation technique, initially developed by Quenouille (1949) and expended and named by Tukey (1958). The basic idea of the jackknife technique is to systematically leave out each observation from a dataset and calculate the parameter estimate over the remaining observations. Then, these calculations are aggregated for specific statistical purposes (Quenouille, 1956; Tukey, 1958). This section shows how to derive the jackknife residual for linearized pseudorange-based positioning systems and develop the hypothesis test to detect potential faults.

(1) Full set solution based on weighted least square

With n measurements, the estimated receiver positioning state $\hat{\mathbf{x}}$ can be solved by the weighted least square (WLS) method (in an iterative approach) as follows:

$$\Delta\hat{\mathbf{x}}{=}\mathbf{S}\mathbf{y} \tag{4a}$$

$$\hat{\mathbf{x}}{=}\mathbf{x}_0 + \Delta\hat{\mathbf{x}} \,, \tag{4b}$$

where $\mathbf{S}$ is the weighted least square solution matrix for a full set of n measurements

$$\mathbf{S} = \left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W} \,, \tag{5}$$

and $\mathbf{W}$ is the weighting matrix and usually takes the inverse of the covariance matrix of $\boldsymbol{\varepsilon}$.

(2) Subset solution based on weighted least square

The linearized pseudorange-based positioning system for the ith subset (i.e., with the ith measurement excluded) can be written by

$$\mathbf{y}^{(i,*)} = \mathbf{H}^{(i,*)}\Delta\mathbf{x}^{(i)} + \boldsymbol{\varepsilon}^{(i,*)} \tag{6}$$

where $\mathbf{y}^{(i,*)}$, $\boldsymbol{\varepsilon}^{(i,*)}$, $\Delta\mathbf{x}^{(i)}$ and $\mathbf{H}^{(i,*)}$ have the same meaning as that in (2) but are defined on the ith subset. Note that $\mathbf{y}^{(i,*)}$ and $\boldsymbol{\varepsilon}^{(i,*)}$ are $n \times 1$ vectors, $\Delta\mathbf{x}^{(i)}$ is an $m \times 1$ vector, and $\mathbf{H}^{(i,*)}$ is a $(n-1) \times 4$ matrix. Define the $m \times (n-1)$ matrix $\mathbf{S}^{(i,*)}$ as the solution matrix on the ith subset

$$\mathbf{S}^{(i,*)} = \left(\mathbf{H}^{(i,*)^T}\mathbf{W}^{(i,*)}\mathbf{H}^{(i,*)}\right)^{-1}\mathbf{H}^{(i,*)^T}\mathbf{W}^{(i,*)} \,, \tag{7}$$

where $\mathbf{W}^{(i,*)}$ is the weight matrix of size of $(n-1)\times(n-1)$. Alternatively, $\mathbf{S}^{(i)}$ can be constructed so that it operates on the full of measurements by extending $\mathbf{S}^{(i,*)}$ with an ith zeroed column,

$$\mathbf{S}^{(i)} = \begin{bmatrix} \mathbf{S}^{(i,*)}_{:,1:i-1} & 0 & \mathbf{S}^{(i,*)}_{:,i:n-1} \end{bmatrix}, \tag{8}$$

where $\mathbf{S}^{(i,*)}_{:,1:i-1}$ is the first $i-1$ columns of $\mathbf{S}^{(i,*)}$ and $\mathbf{S}^{(i,*)}_{:,i:n-1}$ is the remaining columns of $\mathbf{S}^{(i,*)}$. Then, the subsolutions are given by

$$\Delta\hat{\mathbf{x}}^{(i)} = \mathbf{S}^{(i)}\mathbf{y}^{(i)} \;\forall i = 1\cdots n \tag{9a}$$

$$\hat{\mathbf{x}}^{(i)} = \mathbf{x}_0^{(i)} + \Delta\hat{\mathbf{x}}^{(i)} \;\forall i = 1\cdots n\,, \tag{9b}$$

where $\mathbf{x}_0^{(i)}$ is the linearized point of the subsolution, and $\mathbf{y}^{(i)}$ is the measurement vector of the subsolution. Note that $\mathbf{y}^{(i)}$ is also a $n\times 1$ vector, but $\mathbf{y}^{(i)}$ and $\mathbf{y}$ are different. As shown in Equation (3), $\mathbf{y}^{(i)}$ and $\mathbf{y}$ are the function of the raw measurement and the linearized point. Since $\mathbf{y}^{(i)}$ and $\mathbf{y}$ have different linearized points, their values are consequently different.

(3) Construction of Jackknife residual

The predicted ith measurement with the subsolution $\Delta\hat{\mathbf{x}}^{(i)}$ is given by

$$\hat{y}_i = \mathbf{g}_i\Delta\hat{\mathbf{x}}^{(i)} \tag{10a}$$

$$\mathbf{g}_i = \mathbf{h}\big(\{\mathbf{p}^{i,j}\}, \mathbf{x}_0^{(i)}\big)\,, \tag{10b}$$

where $\mathbf{h}$ is defined in Section II.1. The Jackknife residual is given by the difference between y_i and $\hat{y}_i$

$$t_i = y_i - \hat{y}_i\,, \tag{11}$$

where

$$y_i = f(\rho_i, \mathbf{x}_0^{(i)})\,. \tag{12}$$

(4) Distribution of Jackknife residual under Gaussian noises

Based on the subsolution $\Delta\hat{\mathbf{x}}^{(i)}$, the linearized measurement model related to y_i can be written by

$$y_i = \mathbf{g}_i\Delta\mathbf{x}^{(i)} + \varepsilon_i\,. \tag{13}$$

Then

$$\begin{aligned} t_i &= \big(\mathbf{g}_i\Delta\mathbf{x}^{(i)} + \varepsilon_i\big) - \mathbf{g}_i\Delta\hat{\mathbf{x}}^{(i)} \\ &= \mathbf{g}_i\big(\Delta\mathbf{x}^{(i)} - \Delta\hat{\mathbf{x}}^{(i)}\big) + \varepsilon_i\,. \end{aligned} \tag{14}$$

Assume ε_i has a zero-mean Gaussian distribution

$$\varepsilon_i \sim \mathcal{N}\left(0, \sigma_i^2\right) \;\forall i = 1\cdots n\,. \tag{15}$$

The distribution of t_i is given by (a proof is provided in Appendix B)

$$t_i \sim \mathcal{N}\left(0, \mathbf{g}_i\mathbf{S}^{(i)}\mathbf{W}^{-1}{\mathbf{S}^{(i)}}^T\mathbf{g}_i^T + \sigma_i^2\right)\,. \tag{16}$$

(5) Jackknife test for fault detection

Formalize the following hypotheses:

$$\begin{aligned} &H_0^{(i)}\text{: No failure in the } i\text{th measurement} \\ &H_1^{(i)}\text{: A failure exists in the } i\text{th measurement}\,. \end{aligned} \tag{17}$$

Origin test: $H_0^{(i)}$ is rejected if $|t_i| > \left(\mathbf{g}_i\mathbf{S}^{(i)}\mathbf{W}^{-1}\mathbf{S}^{(i)^T}\mathbf{g}_i^T + \sigma_i^2\right)^{\frac{1}{2}} Q^{-1}(\frac{\alpha}{2})$ at significant level of α, where $Q^{-1}(\cdot)$ is the quantile function of a standard normal variable. The probability of type I error (false alarm) of the origin test is α.

In practice, the above test will be conducted for each subsolution to detect the potential failure in measurements, which evolves into a multiple-testing problem. In such a case, the type I error is actually enlarged. Thus, the following hypotheses are formalized instead, which are known as the Bonferroni correction (Bonferroni, 1936):

$$\begin{aligned} &H_0\text{: No failure in the } n \text{ measurements} \\ &H_1\text{: At least one failure exists in the } n \text{ measurements}\,. \end{aligned} \tag{18}$$

Corrected test: H_0 is rejected if $|t_i| > \left(\mathbf{g}_i\mathbf{S}^{(i)}\mathbf{W}^{-1}\mathbf{S}^{(i)^T}\mathbf{g}_i^T + \sigma_i^2\right)^{\frac{1}{2}} Q^{-1}(\frac{\alpha}{2n})$ at significant level of α^*. The probability of type I error (false alarm) of the corrected test is α^*.

In implementing the corrected test, an upper limit τ (e.g., 0.05) will be set for α^*. Then the type I error α of the individual test would be $\frac{\tau}{n}$ (as shown in Appendix C), which could be very small when n takes a large value. Therefore, the individual test and the corrected test both could be conservative. However, in satellite navigation applications, it is rare to a large n, which ensures the feasibility of the corrected test.

(6) Notes on multiple hypothesis solution separation

The separation of the vertical position between the full solution and the ith subsolution is given by (Blanch et al., 2010; Brenner, 1996)

$$\begin{aligned} d_i=&\hat{\mathbf{x}}_v - \hat{\mathbf{x}}_v^{(i)} \\ =&\mathbf{x}_{0,v} - \mathbf{x}_{0,v}^{(i)} + \Delta\hat{\mathbf{x}}_v - \Delta\hat{\mathbf{x}}_v^{(i)}\,, \end{aligned} \tag{19}$$

where the subscript v represents the vertical component of the positioning solution. Assuming that ε_i takes the same form in Equation (15), d_i will have a Gaussian distribution

$$d_i \sim \mathcal{N}\Big(0, \big(\mathbf{s}_3 - \mathbf{s}_3^{(i)}\big)\mathbf{W}^{-1}\big(\mathbf{s}_3 - \mathbf{s}_3^{(i)}\big)^T\Big)\,, \tag{20}$$

where $\mathbf{s}_3$ and $\mathbf{s}_3^{(i)}$ are the 3rd row (the vertical axis) of $\mathbf{S}$ and $\mathbf{S}^{(i)}$, respectively. For single fault detection, the corrected test in Equation (18) can also be applied for multiple hypothesis solution separation (MHSS), and the threshold is determined based on the distribution of d_i. Applications can be found in (Blanch et al., 2010, 2015; Brenner, 1996).

3. Extend Jackknife Test Detector to Non-Gaussian Noises

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

$$\hat{\mathbf{y}}^{(i,*)} = \mathbf{H}^{(i,*)}\Delta\hat{\mathbf{x}}^{(i)}\,, \tag{21}$$

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

$$\tilde{\mathbf{y}}^{(i)} = \begin{bmatrix} \hat{\mathbf{y}}_{1:i-1,:}^{(i,*)} \\ \hat{y}_i \\ \hat{\mathbf{y}}_{i:n-1,:}^{(i,*)} \end{bmatrix}, \tilde{\mathbf{H}}^{(i)} = \begin{bmatrix} \mathbf{H}_{1:i-1,:}^{(i,*)} \\ \mathbf{g}_i \\ \mathbf{H}_{i:n-1,:}^{(i,*)} \end{bmatrix}\,, \tag{22}$$

where $\hat{\mathbf{y}}_{1:i-1,:}^{(i,*)}$ is the first $i-1$ rows of $\hat{\mathbf{y}}^{(i,*)}$ and $\hat{\mathbf{y}}_{i:n-1,:}^{(i,*)}$ is the remaining rows of $\hat{\mathbf{y}}^{(i,*)}$. A similar definition is given for $\mathbf{H}_{1:i-1,:}^{(i,*)}$ and $\mathbf{H}_{i:n-1,:}^{(i,*)}$. The measurement residual is

$$\begin{aligned} \mathbf{y}^{(i)} - \tilde{\mathbf{y}}^{(i)} =& \mathbf{y}^{(i)} - \tilde{\mathbf{H}}^{(i)}\Delta\hat{\mathbf{x}}^{(i)} \\ =& \left(\mathbf{I} - \tilde{\mathbf{P}}\right)\mathbf{y}^{(i)}\,, \end{aligned} \tag{23}$$

where

$$\tilde{\mathbf{P}} = \tilde{\mathbf{H}}^{(i)}\mathbf{S}^{(i)}\,. \tag{24}$$

According to Equations (2) and (6), $\mathbf{y}^{(i)}$ can be re-written by

$$\mathbf{y}^{(i)} = \tilde{\mathbf{H}}^{(i)}\Delta\mathbf{x}^{(i)} + \boldsymbol{\varepsilon}\,, \tag{25}$$

where $\boldsymbol{\varepsilon}$ is the measurement noise vector with arbitrary distributions. Then, Equation (23) can be written by

$$\mathbf{y}^{(i)} - \tilde{\mathbf{y}}^{(i)} = \left(\mathbf{I} - \tilde{\mathbf{P}}\right)\tilde{\mathbf{H}}^{(i)}\Delta\mathbf{x}^{(i)} + \left(\mathbf{I} - \tilde{\mathbf{P}}\right)\boldsymbol{\varepsilon}\,. \tag{26}$$

It is proved in Appendix D that

$$\left(\mathbf{I} - \tilde{\mathbf{P}}\right)\tilde{\mathbf{H}}^{(i)} = 0\,. \tag{27}$$

Therefore,

$$\mathbf{y}^{(i)} - \tilde{\mathbf{y}}^{(i)} = \left(\mathbf{I} - \tilde{\mathbf{P}}\right)\boldsymbol{\varepsilon}\,. \tag{28}$$

Define $\tilde{\mathbf{p}}_i$ as the ith row of $\left(\mathbf{I} - \tilde{\mathbf{P}}\right)$, then the Jackknife residual is given by

$$t_i = \tilde{\mathbf{p}}_i\boldsymbol{\varepsilon}\,. \tag{29}$$

Equation (29) can be rewritten as the linear combination of measurement noises as follows:

$$t_i = \sum_{j=1}^{n}\tilde{p}_{i,j}\varepsilon_j\,, \tag{30}$$

where $\tilde{p}_{i,j}$ is the jth element of $\tilde{\mathbf{p}}_i$. Remarkably, ε_j can have an **arbitrary distribution** as long as it has a probability density function (PDF) $f_{\varepsilon_j}(\cdot)$. Then, the PDF of t_i can be derived as (a proof is given in Appendix E)

$$f_{t_i}(x) = \prod_{j=1}^{n}|\tilde{p}_{i,j}|^{-1} f_{\varepsilon_1}\left(\frac{x}{|\tilde{p}_{i,1}|}\right) * f_{\varepsilon_2}\left(\frac{x}{|\tilde{p}_{i,2}|}\right) * \ldots * f_{\varepsilon_n}\left(\frac{x}{|\tilde{p}_{i,n}|}\right)\,, \tag{31}$$

where $*$ denotes the convolution operation. Since the distribution of t_i is known, the jackknife test illustrated in Section II.2 can be applied under both Gaussian and non-Gaussian noises. Moreover, it is proved in Appendix B that the construction process of the Jackknife residual in this section is equivalent to that in Section II.2 under Gaussian noise settings.

III. OVERBOUNDING NON-GAUSSIAN NOISES

To characterize and simplify the error profile of measurements in the above pseudorange-based positioning systems, the concept of overbound is proposed and widely studied in the navigation communities (Braff & Shively, 2005; Larson et al., 2019; Shively, 2001). Overbound represents the worst possible error distribution in the absence of a hardware fault (DeCleene, 2000; Rife, Pullen, & Pervan, 2004). Researchers have proposed various Gaussian, semi-Gaussian, and non-Gaussian overbounding methods. A review of these methods can refer to Rife and Pervan (2012). This section will introduce two practical overbounding methods, including the Gaussian overbound DeCleene (2000) and the Principal Gaussian overbound (Yan, Zhong, & Hsu, 2024). The latter one is a non-Gaussian overbounding method. Both methods will be used for bounding zero-mean non-Gaussian measurement error in Sections IV-VI.

1. Gaussian overbound

Let the CDF of the random variable v be G_v. The Gaussian overbound is determined by finding the minimum δ that satisfies

$$\int_{-\infty}^{x} f_{\mathcal{N}}\left(x;0,\delta\right)dx \geq G_v(x)\ \forall x < 0 \tag{32a}$$

$$\int_{-\infty}^{x} f_{\mathcal{N}}\left(x;0,\delta\right)dx \leq G_v(x)\ \forall x \geq 0\,, \tag{32b}$$

where $f_{\mathcal{N}}(x;0,\sigma)$ is the PDF of a zero-mean Gaussian distribution with a standard deviation of σ.

2. Principal Gaussian overbound

The Principal Gaussian overbound (Yan, Zhong, & Hsu, 2024) utilizes the zero-mean bimodal Gaussian mixture model (BGMM) to fit the error distribution based on the expectation–maximization (EM) algorithm (Dempster et al., 1977) and divides the BGMM into the core and tail regions based on the analysis of BGMM membership weight. Within each region, one of the Gaussian components in the BGMM holds a dominant position, and a CDF overbound is constructed based on the dominant Gaussian component. The PDF of the Principal Gaussian overbound (PGO) is given by

$$f_{PGO}(x) = \begin{cases} (1+k)\,(1-p_1)\,f_{\mathcal{N}}\left(x;0,\delta_2^2\right) & |x| > x_{rp} \\ p_1 f_{\mathcal{N}}\left(x;0,\delta_1^2\right) + c & |x| \le x_{rp} \end{cases}, \tag{33}$$

where $f_{\mathcal{N}}\left(x;0,\delta_1^2\right)$ and $f_{\mathcal{N}}\left(x;0,\delta_2^2\right)$ are the PDF of the first and the second Gaussian component of the fitted BGMM, δ_1 and δ_2 are the corresponding standard deviations, and p_1 and $1-p_1$ are the mixing weight of the two Gaussian components, respectively; k, c, and x_{rp} are parameters uniquely determined by the partition strategy based on the analysis of BGMM membership weight, as shown in Appendix F. The variance of PGO can be calculated by

$$\mathrm{Var}(x) = \int_{-\infty}^{\infty} x^2 f_{PGO}(x)dx\,. \tag{34}$$

A detailed description of PGO can refer to Yan, Zhong, and Hsu (2024) and is also presented in Appendix F. Soon, it will be shown in Sections IV-VI that PGO provides a sharper yet conservative overbound than the Gaussian overbound for heavy-tailed error distribution. Notably, it is proved that PGO can maintain the overbounding property through convolution (Yan, Zhong, & Hsu, 2024), which is the basis for deriving pseudorange-level requirements from the position domain integrity requirements (DeCleene, 2000).

IV. DETECTION PERFORMANCE WITH WORLDWIDE SIMULATIONS

This section shows the single-failure detection results of a set of users distributed over the world during one day. MAAST (Jan et al., 2001), a MATLAB toolset developed at Stanford University, is utilized to simulate the dual frequency pseudorange measurements, satellite positions, and user locations. Specifically, the 27-satellite Galileo constellation (Zandbergen et al., 2004) is used to simulate satellite positions. The users are placed on a grid every 10 degrees longitude and latitude (which gives 648 locations). For each location, the geometries are simulated every 5 min (which gives 288 time steps). Two types of measurement noises for dual frequency pseudorange measurements are simulated, including the Gaussian distribution and the Normal Inverse Gaussian (NIG) distribution. The dual pseudorange frequency measurements are then simulated by adding the randomly generated sample from the given error distribution to the true range. The measurement model is given by dual-frequency SPP, as shown in Appendix A.1. For each time and user location, an artificial bias ($10\,\mathrm{m}$) is injected into one of the measurements, and the detection rate is defined as

$$P_{dec} = \frac{\text{Detected epochs in one day}}{\text{Valid epochs in one day}}, \tag{35}$$

where the nominator could be less than 288 since the number of satellites in view may not satisfy the minimum requirements for fault detection (for SPP, the minimum number is 5).

1. Simulation with Gaussian Distributed Measurement Errors

The measurement error ε_i is given by

$$\varepsilon_i \sim \mathcal{N}(0,\sigma_i^2)\,, \tag{36}$$

where

$$\sigma_i^2 = \left(\sigma_{eph\&clk,i}^{L1\text{-}L5}\right)^2 + \sigma_{trop,i}^2 + \left(\frac{\gamma}{\gamma-1}\right)^2\left(\sigma_{cnmp,i}^{L1}\right)^2 + \left(\frac{1}{\gamma-1}\right)^2\left(\sigma_{cnmp,i}^{L5}\right)^2, \gamma = \left(\frac{f_{L1}}{f_{L5}}\right)^2, \tag{37}$$

$\sigma_{eph\&clk,i}^{L1\text{-}L5}$,$\sigma_{trop,i}$, $\sigma_{cnmp,i}^{L1}$, $\sigma_{cnmp,i}^{L5}$ are the standard deviation of the satellite clock and ephemeris error for the L1-L5 ionospheric free combination, tropospheric delay and L1 and L5 code noise and multipath errors, respectively; and f_{L1} and f_{L5} are the L1 and L5 frequencies, respectively. The configuration of these error terms is described in Appendix G.

The standard jackknife test (JT) detector in Equations (11)-(16) and the multiple hypothesis solution separation (MHSS) detector in Equations (19)-(20) are deployed to detect the faults for each time and user location. The weighting matrix $\mathbf{W}$ used in the two detectors takes the same value, which is the inverse of the covariance matrix of the measurement error. Figure 1 shows the

detection rate contour of the two detectors. As can be seen, the JT detector demonstrates identical performance as the MHSS detector, where both detectors show more than $99.5\,\%$ detection rate at most user locations.

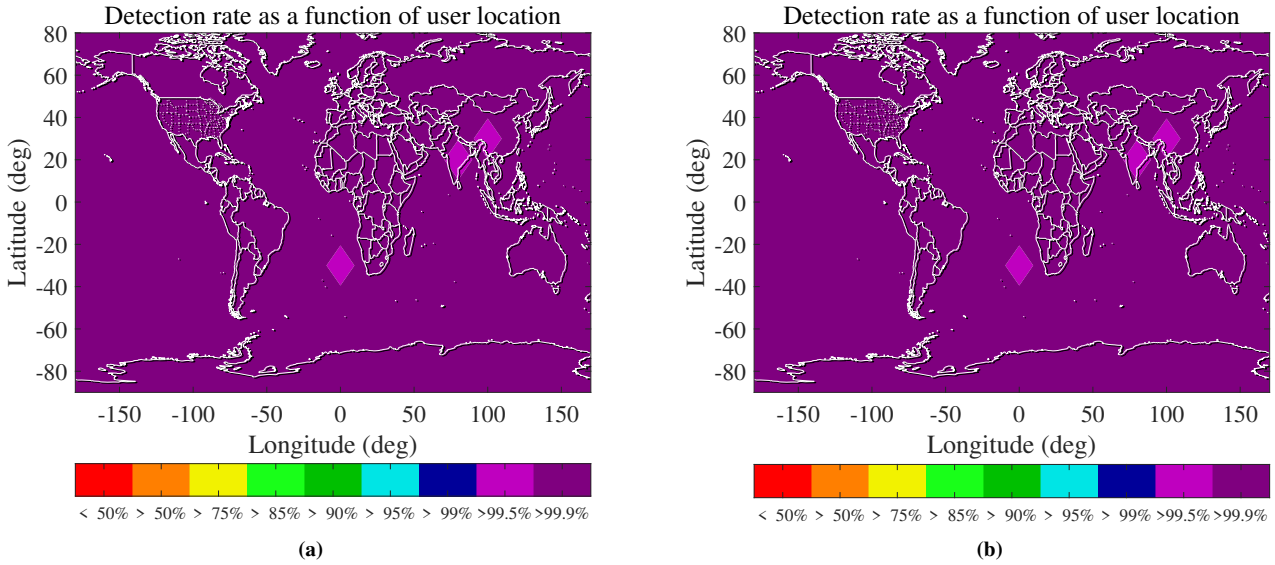


Figure 1: Detection performance of (a) the MHSS and (b) the JT detectors with artificially injected bias ($10\,\mathrm{m}$) under Gaussian distributed measurement errors.

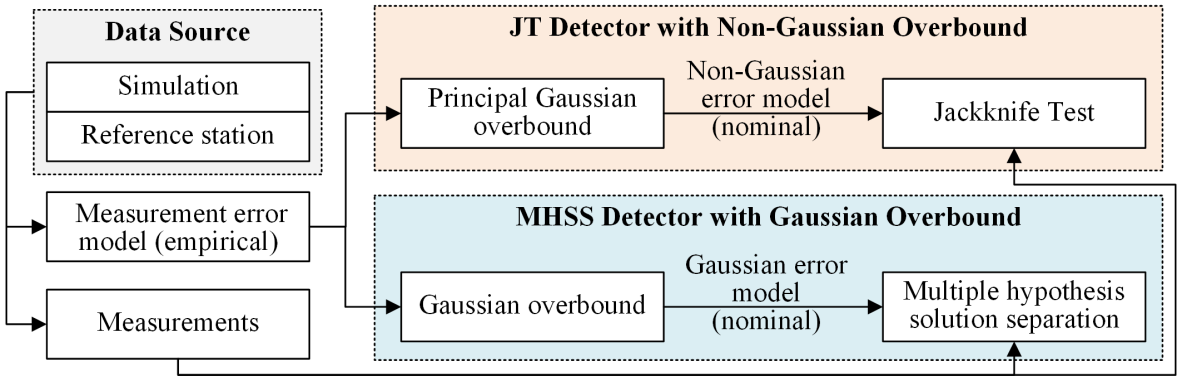


Figure 2: The flowchart of the fault detection process in the MHSS and the JT detectors. The empirical measurement error model is obtained from either simulation or the collected data in reference stations. Since the empirical measurement error distribution could be arbitrary, overbounding methods are employed to obtain the simplified nominal error model for fault detection purposes. The Gaussian overbound is employed for the MHSS detector, while the Principal Gaussian overbound is employed for the JT detector.

2. Simulation with NIG Distributed Measurement Errors

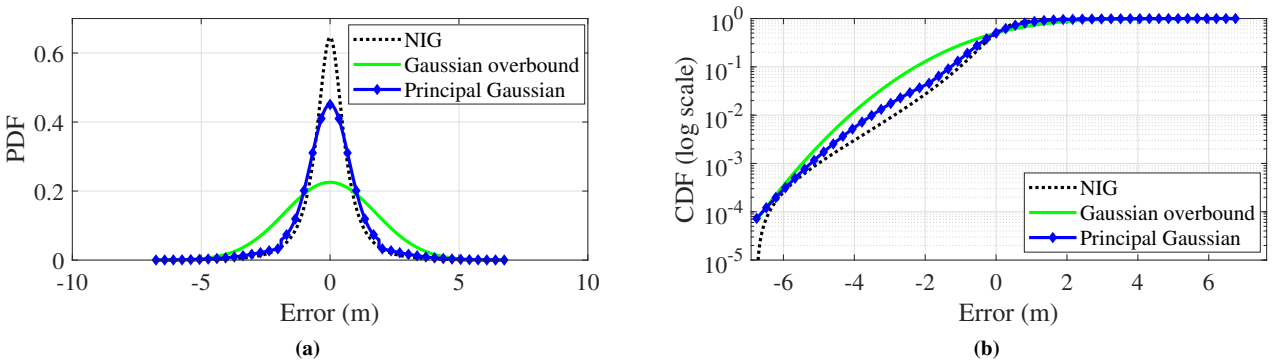
The NIG distribution is found to be useful in approximating the heavy tails of the ground-station error distribution for LAAS (Braff & Shively, 2005; Rife, Pullen, & Pervan, 2004), and its PDF is given by

$$f_{NIG}(x) = \frac{\delta_0^2 \exp\left(\delta_0^2\right)}{\pi\sqrt{x^2+\delta_0^2}} K_1\left(\delta_0\sqrt{x^2+\delta_0^2}\right), \tag{38}$$

where δ_0 is the shape parameter that determines the weight of the NIG tail, and K_1 is a modified Bessel function of the second kind, degree one. Similar to the setting in Braff and Shively (2005), $\delta_0 = 0.65$ is adopted to simulate the measurement error ε_i. Note that the error distribution for each measurement is set to be identical, which is different from the Gaussian case in IV.1. The primary reason for this simplification is that, to the best of the authors' knowledge, there is no developed method

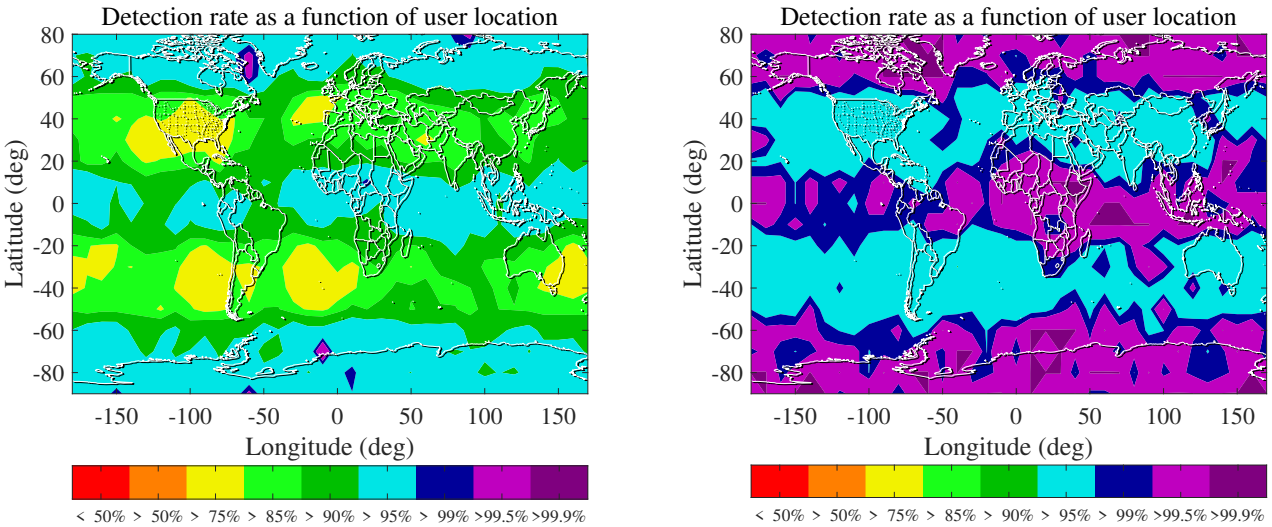
for establishing the geometry-related NIG error model. Since the simplification is applied for both detectors, its impact on the conclusion drawn from the subsequent experimental results is expected to be negligible.

The nominal measurement error is obtained by bounding the NIG distributed measurement error and used in fault detection algorithms. The MHSS and the extended JT detectors adopt different overbounding methods, as shown in Figure 2. Specifically, the Gaussian overbound (DeCleene, 2000) introduced in III.1 is employed to bound the NIG distributed measurement error, where the variance of Gaussian overbound is used for constructing the weighting matrix W_{MHSS} (a diagonal matrix) in the MHSS detector. For the extended JT detector in Equations (29)-(31), which is capable of handling non-Gaussian conditions, the Principal Gaussian overbound (PGO) illustrated in III.2 is adopted instead. The weighting matrix W_{JT} for the extended JT detector is also a diagonal matrix, with each diagonal element taking the inverse of the variance of PGO, as shown in Equation (34). Figure 3 plots the Gaussian overbound and the PGO of the NIG distribution ($\delta_0 = 0.65$). As can be seen, PGO provides a sharper bound for the NIG distribution at both the tail and core regions than the Gaussian overbound.



(a) **(b)**

Figure 3: The (a) PDF and (b) CDF (plotted in logarithmic scale) of the Gaussian overbound and Principal Gaussian overbound for the NIG distribution ($\delta_0 = 0.65$).



(a) **(b)**

Figure 4: Detection performance of (a) the MHSS and (b) the JT detectors with artificially injected bias (10 m) under NIG distributed measurement errors.

Figure 4 shows the contour plot of the detection rate for both detectors, revealing a noticeable disparity. The JT detector exhibits a substantial enhancement in detection rate when compared to the MHSS detector. In most user locations, the JT detector achieves a detection rate of over 95 %. Moreover, in a considerable number of user locations, the detection rate even surpasses 99.5 %. However, the MHSS detector exhibits a detection rate of 85 % at nearly half of the locations. The superiority of the

JT detector can be primarily attributed to two factors: 1) a sharper overbound for non-Gaussian measurement error and 2) the accommodation of the jackknife test for non-Gaussian noises. On the one hand, the sharper overbound provided by PGO can better characterize the measurement error distribution than the Gaussian overbound, as shown in Figure 3, providing an accurate probabilistic model for hypothesis testing. It is known that an inaccurate probabilistic model can hinder the performance of hypothesis testing, such as inflating the type I error (false alarm) or type II error (miss detection), and even make the testing invalid. On the other hand, the theoretical underpinnings of the jackknife test provide a solid basis for its application in fault detection with non-Gaussian noises (or say, non-Gaussian overbounds). The threshold for the JT detector is derived through rigorous mathematical derivations rather than relying on simulation techniques. These two factors complement each other, resulting in the exceptional performance of the JT detector.

V. DETECTION PERFORMANCE FOR CORS DGNSS POSITIONING

The analysis conducted in IV.2 primarily relies on simulated data, targeting a specific non-Gaussian measurement error model known as NIG. To showcase the practical applicability of the JT detector, this section focuses on evaluating its performance in the context of a DGNSS positioning with data collected from two Continuously Operating Reference Stations (CORS) in Minneapolis. A brief introduction to DGNSS error and positioning is presented in Appendix A.2.

1. Data Prepossessing and Overbounding

This section utilizes reference station data from the CORS website, operated by the National Geodetic Survey (NGS). Specifically, data is collected from stations MNAV and ZMP1, which are approximately 11.5 kilometers apart. The positions of the two reference stations are obtained from the RINEX file header, and the satellite position is calculated based on the broadcast ephemeris from NASA's Archive of Space Geodesy Data website (NASA, n.d.) by utilizing RTKLIB (Takasu & Yasuda, 2009). For each double-differenced pseudorange in Equation (49), the satellite with the highest elevation angle at that specific time epoch is assigned as the master satellite. The research encompasses data from January 1st, 2020, to January 31st, 2020, with a focus solely on GPS data. Since GPS data are strongly influenced by the elevation angle, the collected data are organized into bins based on elevation angles (particularly refer to the elevation angle between the user receiver and the slave satellite) every 5° from 15° to 80°, encompassing the highest observed elevation angle in the dataset.

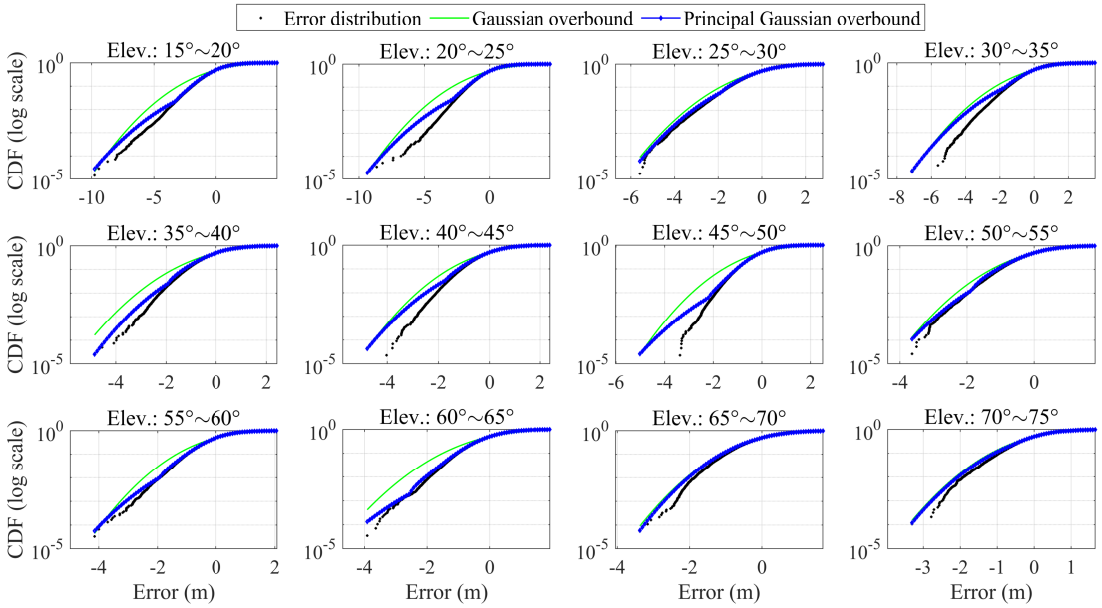


Figure 5: The CDF (plotted in logarithm scale) of the Gaussian overbound and Principal Gaussian overbound for DGNSS errors in each elevation angle bin.

Within each bin, the Gaussian overbound and the PGO for DGNSS errors are computed and plotted in Figure 5. In low elevation angle bins (15° to 35°), the PGO yields a significantly sharper bound than the Gaussian overbound. As discussed in (Yan, Zhong, & Hsu, 2024), DGNSS error in low elevation angles has heavy tails, which is the main target of the PGO, thereby the PGO exhibits better performance than the Gaussian overbound. In middle elevation angle bins (35° to 65°), the PGO has a slight superiority in bounding DGNSS errors to the Gaussian overbound, while the bounding performance of the two methods

is nearly the same in high elevation angle bins ($> 65^\circ$). Similar results are found in the complementary cumulative distribution function (CCDF) plot of the Gaussian overbound and the PGO for DGNSS errors, which are presented in the Appendix H.

2. Detection Performance with Artificially Injected Bias

Within each elevation angle bin, the variance of the overbounding distribution is calculated and is taken as the variance of measurement errors for all DGNSS measurements whose elevation angle falls in the bin. Similar to Section IV.2 (as suggested in Figure 2), PGO is used to calculate the variance and is taken as the nominal error model in the JT detector, while the Gaussian overbound is utilized in the MHSS detector. For example, in constructing the JT detector, the DGNSS measurements with elevation angles ranging from 30° to 35° have the same variance of measurement errors, which is the variance of PGO within that bin. Indeed, this simplification unavoidably hinders the accuracy of error modeling. A more rigorous way is to characterize the error profile for each discrete elevation angle (could be an infinite number of discrete values); however, such the characterization process is impractical with limited data and is out of the scope of our research.

The detecting performance of the JT detector and the MHSS detector is examined on data collected on January 1st, 2020. The GPS pseudorange data are retrieved every 30 seconds, resulting in 2,880 epochs in total. At each epoch, an artificial bias is injected into one of the measurements. Ten experiments are performed individually, each involving the injection of bias with a specific magnitude. The detection rates of both detectors are tabulated in Table 1. For all ten cases, the JT detector significantly outperforms the MHSS detector in terms of detection rate. When the injected bias falls within the range of $3\,\mathrm{m}$ to $6\,\mathrm{m}$, the detection rate of the JT detector is over $13\,\%$ higher than that of the MHSS detector, demonstrating the superior of the JT detector in detecting small faults. Remarkably, when the injected bias is set at $4\,\mathrm{m}$, the JT detector achieves the detection rate of $60.92\,\%$, surpassing that of the MHSS detector by $20.13\,\%$. When the magnitude of the injected bias continues to increase, both detectors yield satisfactory detection performance; however, the JT detector still performs better than the MHSS detector.

Table 1: The detection rate of the JT and the MHSS detectors on the CORS DGNSS data with different magnitudes of the injected bias.

Inject bias	MHSS[1]	JT[2]	Improvement	Inject bias	MHSS	JT	Improvement
1 m	0.31 %	**2.30** %	1.99 %	6 m	71.93 %	**84.95** %	13.03 %
2 m	3.31 %	**12.68** %	9.37 %	7 m	81.71 %	**92.51** %	10.80 %
3 m	18.57 %	**37.76** %	19.19 %	8 m	89.62 %	**96.55** %	6.93 %
4 m	40.79 %	**60.92** %	20.13 %	9 m	94.74 %	**98.71** %	3.97 %
5 m	59.42 %	**74.57** %	15.15 %	10 m	96.90 %	**99.48** %	2.58 %

[1] "MHSS" is short for "multiple hypothesis solution separation"
[2] "JT" is short for "jackknife test"

VI. APPLICATION TO REAL-WORLD SATELLITE CLOCK ANOMALY DETETION

This section applies the proposed method to real-world satellite clock anomaly detection in the context of single point positioning (SPP). A brief introduction to SPP is given in Appendix A.1. GPS PRN-1 experienced a clock anomaly on January 28th, 2023, where the anomaly began at GPS time 15:02:30 and was set unhealthy at 16:05:00 (Lai et al., 2023). The clock anomaly began to be corrected by manual control at 18:00, and the correction was completed at 20:00:00. The anomaly happened when PRN-1 was over in the middle of South Pacific, so the data collected from the nearby station, CHTI, located in Chatham Island, New Zealand, are used for analysis. Specifically, the observation data at CHTI from January 1st, 2023, to January 31st, 2023 are collected from the CORS website. The same method used in V.1 is applied to determine the position of CHTI and satellites and to overbound the pseudorange measurement error (only L1 measurements are studied, and faulty measurements are pre-excluded). Notably, the pseudorange measurement is corrected by clock bias correction, ionospheric corrections, and tropospheric corrections based on RTKLIB (Takasu & Yasuda, 2009) with the broadcast ephemeris from NASA's Archive of Space Geodesy Data website (NASA, n.d.). The PGO and Gaussian overbounding results are listed in Appendix I.

Figure 6 shows the detection states of the MHSS and JT detectors at the CHTI station on January 28th. Similar to Section IV.2 (as suggested in Figure 2), PGO is used to construct the nominal error model in the JT detector while the Gaussian overbound is adopted for the MHSS detector. PRN-1 is visible for CHTI in two time periods, i.e., 05:44:30∼08:13:30 and 14:08:30∼18:23:00, as marked by the green shade area in Figure 6. In addition, the fault period 15:02:30∼20:00:00 is marked by the red shade area. As can be seen, the JT detector can detect the anomaly at most times of the fault period, while the MHSS detector occasionally declares faults during that period. Remarkably, the JT detector claims the fault at 15:03:00, merely 30 seconds after the anomaly happened, which is 8 minutes earlier than the MHSS detector. Since the sampling rate of the observation data retrieved from the CORS website is 30 seconds, the minimum delayed time of detection for the JT detector is

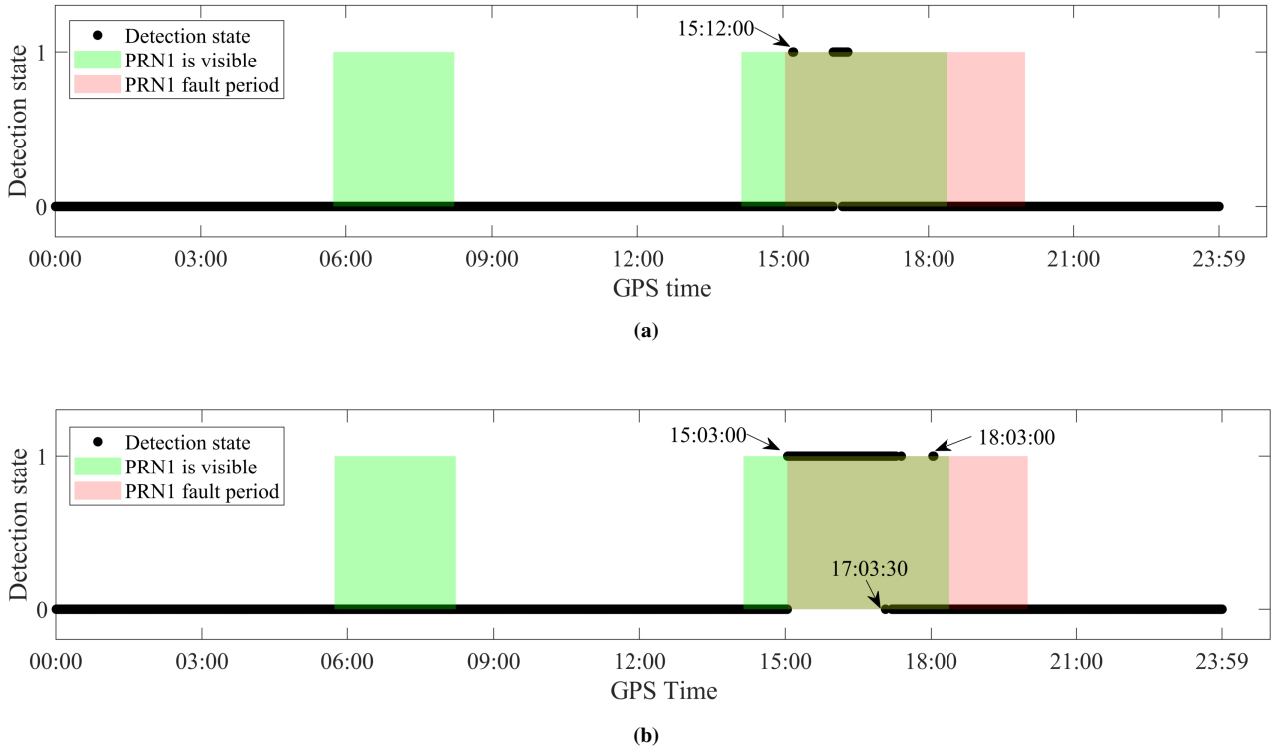


Figure 6: Detection states of (a) the MHSS and (b) the JT detectors at the CHTI station on January 28th, 2023. The detection state of "1" refers to "a fault is claimed" while "0" refers to "not failure is claimed".

limited to 30 seconds. With higher frequency observation data, the JT detector could yield a smaller delayed time of detection.

It is worth noting that a considerably long period of miss detection exists when using the JT detector for anomaly detection (not to mention the case of the MHSS detector). The miss-detection period began at 17:03:30 and lasted until 18:02:30, during which the JT detector occasionally claims faults. The primary reason can be attributed to the inaccuracy of the nominal error model. Since only one-month data is used to model the pseudorange measurement error and compute the overbounds, the nominal error model could be conservative and can even shield small faults. Therefore, the detection performance of the JT detector could be degraded. Nevertheless, the JT detector still shows significantly better performance than the MHSS detector.

VII. CONCLUSIONS AND FUTURE WORK

This paper proposes a fault detection method with rigorous derivations for linearized pseudorange-based positioning systems under non-Gaussian noises. This paper contributes to the literature mainly in two aspects. First, this paper develops the jackknife test (JT) detector by constructing a test statistic with rigorous statistic properties, providing the theoretical basis to detect faulty measurements in linearized pseudorange-based positioning systems under non-Gaussian noises. Second, this paper demonstrates the outstanding performance of the JT detector on fault detection tasks in the single-frequency SPP, dual-frequency SPP, and DGNSS positioning systems, respectively, through both simulated and real-world experiments, which exemplifies the feasibility and wide adaptability of the proposed method.

This work provides the theoretical foundation for detecting faults in linearized pseudorange-based positioning systems under non-Gaussian noises. This, indeed, provides the necessary theoretical basis to develop multi-hypothesis integrity monitoring algorithms. Since the jackknife test statistic is developed in the measurement domain, efforts are anticipated to establish the relationship between the jackknife test statistic and the integrity of positioning solutions, which will be the main focus of our future work. In addition, the proposed method is designed for single-fault detection. For multiple-fault detection and integrity monitoring under non-Gaussian noises, a possible solution could be examining the most influential measurements with multiple jackknife tests, which also remain for further exploration.

REFERENCES

Aitken, A. C. (1936). On least squares and linear combination of observations. *Proceedings of the Royal Society of Edinburgh*, *55*, 42–48.

Blanch, J., Walker, T., Enge, P., Lee, Y., Pervan, B., Rippl, M., Spletter, A., & Kropp, V. (2015). Baseline advanced RAIM user algorithm and possible improvements. *IEEE Transactions on Aerospace and Electronic Systems*, *51*(1), 713–732.

Blanch, J., Walter, T., & 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., & Lai, F. (2023). Evaluation of tropospheric and mixed single-dual frequency error models in advanced RAIM. *Proceedings of the 2023 International Technical Meeting of The Institute of Navigation*, 892–902.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, *8*, 3–62.

Braff, R., & Shively, C. (2005). A method of over bounding ground based augmentation system (GBAS) heavy tail error distributions. *The Journal of Navigation*, *58*(1), 83–103.

Brenner, M. (1996). Integrated GPS/inertial fault detection availability. *NAVIGATION*, *43*(2), 111–130.

Brown, R. G., & McBurney, P. W. (1988). Self-contained GPS integrity check using maximum solution separation. *NAVIGATION*, *35*(1), 41–53.

DeCleene, B. (2000). Defining pseudorange integrity - overbounding. *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, 1916–1924.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22.

Gabela, J., Kealy, A., Hedley, M., & Moran, B. (2021). Case study of Bayesian RAIM algorithm integrated with spatial feature constraint and fault detection and exclusion algorithms for multi-sensor positioning. *NAVIGATION*, *68*(2), 333–351.

Gao, Z., Ding, S. X., & Cecati, C. (2015). Real-time fault diagnosis and fault-tolerant control. *IEEE Transactions on Industrial Electronics*, *62*(6), 3752–3756.

Gupta, S., Mohanty, A., & Gao, G. (2022). Getting the best of particle and Kalman filters: GNSS sensor fusion using Rao-Blackwellized particle filter. *Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022)*, 1610–1623.

He, P., Liu, G., Tan, C., & Lu, Y.-e. (2016). Nonlinear fault detection threshold optimization method for RAIM algorithm using a heuristic approach. *GPS Solutions*, *20*, 863–875.

Hsu, L. T., Tokura, H., Kubo, N., Gu, Y., & Kamijo, S. (2017). Multiple faulty GNSS measurement exclusion based on consistency check in urban canyons. *IEEE Sensors Journal*, *17*(6), 1909–1917.

Jan, S.-S., Chan, W., Walter, T., & Enge, P. (2001). Matlab simulation toolset for SBAS availability analysis. *Proceedings of the 14th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2001)*, 2366–2375.

Joerger, M., Chan, F.-C., & Pervan, B. (2014). Solution separation versus residual-based RAIM. *NAVIGATION*, *61*(4), 273–291.

Joerger, M., & Pervan, B. (2013). Kalman filter-based integrity monitoring against sensor faults. *Journal of Guidance, Control, and Dynamics*, *36*(2), 349–361.

Joerger, M., & 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.

Lai, Y.-F., Blanch, J., Walter, T., Kahr, E., Leahy, E., Silva, P., & Ellum, C. (2023). Prototyping integrity monitors for PPP fault detections. *Proceedings of the 36th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2023)*, 2592–2605.

Larson, J. D., Gebre-Egziabher, D., & Rife, J. H. (2019). Gaussian-Pareto overbounding of DGNSS pseudoranges from CORS. *NAVIGATION*, *66*(1), 139–150.

Lee, Y. C. (1986). Analysis of range and position comparison methods as a means to provide GPS integrity in the user receiver. *Proceedings of the 42nd Annual Meeting of The Institute of Navigation (1986)*, 1–4.

Liu, Y., & Morton, Y. J. (2022). Improved automatic detection of GPS satellite oscillator anomaly using a machine learning algorithm. *NAVIGATION*, *69*(1). https://doi.org/10.33012/navi.500

Misra, P., & Enge, P. (2006). *Global Positioning System: Signals, Measurements, and Performance* (2. ed). Ganga-Jamuna Press.

NASA. (n.d.). *Broadcast ephemeris data*. Retrieved January 18, 2024, from https://cddis.nasa.gov/Data_and_Derived_Products/GNSS/broadcast_ephemeris_data.html

Niu, X., Chen, Q., Zhang, Q., Zhang, H., Niu, J., Chen, K., Shi, C., & Liu, J. (2014). Using Allan variance to analyze the error characteristics of GNSS positioning. *GPS Solutions*, *18*(2), 231–242.

Oehler, V., Trautenberg, H. L., Luongo, F., Boyero, J.-P., & Lobert, B. (2004). User integrity risk calculation at the alert limit without fixed allocations. *Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2004)*, 1645–1652.

Osechas, O., Misra, P., & Rife, J. (2012). Carrier-phase acceleration RAIM for GNSS satellite clock fault detection. *NAVIGATION*, *59*(3), 221–235.

Pervan, B., Pullen, S., & Sayim, I. (2000). Sigma estimation, inflation, and monitoring in the LAAS ground system. *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, 1234–1244.

Pervan, B. S., Lawrence, D. G., Cohen, C. E., & Parkinson, B. W. (1996). Parity space methods for autonomous fault detection and exclusion using GPS carrier phase. *Proceedings of Position, Location and Navigation Symposium-PLANS'96*, 649–656.

Pervan, B. S., Pullen, S. P., & Christie, J. R. (1998). A multiple hypothesis approach to satellite navigation integrity. *NAVIGATION*, *45*(1), 61–71.

Pfeifer, T., Lange, S., & Protzel, P. (2021). Advancing mixture models for least squares optimization. *IEEE Robotics and Automation Letters*, *6*(2), 3941–3948.

Pfeifer, T., & Protzel, P. (2019). Expectation-maximization for adaptive mixture models in graph optimization. *2019 International Conference on Robotics and Automation (ICRA)*, 3151–3157.

Quenouille, M. H. (1949). Problems in plane sampling. *The Annals of Mathematical Statistics*, *20*(3), 355–375.

Quenouille, M. H. (1956). Notes on bias in estimation. *Biometrika*, *43*(3/4), 353–360.

Ray, T. N., Pierce, J. D., & Bevly, D. M. (2018). A comparison of particle propagation and weight update methods for indoor positioning systems. *Proceedings of the 31st International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2018)*, 3398–3408.

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

Rife, J., & Pervan, B. (2012). Overbounding revisited: Discrete error-distribution modeling for safety-critical GPS navigation. *IEEE Transactions on Aerospace and Electronic Systems*, *48*(2), 1537–1551.

Rife, J., Pullen, S., & Pervan, B. (2004). Core overbounding and its implications for LAAS integrity. *Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2004)*, 2810–2821.

Rosihan, R., Indriyatmoko, A., Chun, S., Won, D., Lee, Y., Kang, T., Kim, J., & Jun, H. (2006). Particle filtering approach to fault detection and isolation for GPS integrity monitoring. *Proceedings of the 19th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2006)*, 873–881.

Shively, C. A. (2001). A comparison of LAAS error bounding concepts. *Proceedings of the 2001 National Technical Meeting of The Institute of Navigation*, 501–511.

Sturza, M. A. (1988). Navigation system integrity monitoring using redundant measurements. *NAVIGATION*, *35*(4), 483–501.

Sünderhauf, N., Obst, M., Lange, S., Wanielik, G., & Protzel, P. (2013). Switchable constraints and incremental smoothing for online mitigation of non-line-of-sight and multipath effects. *2013 IEEE Intelligent Vehicles Symposium (IV)*, 262–268.

Takasu, T., & Yasuda, A. (2009). Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB. *International symposium on GPS/GNSS*, *1*, 1–6.

Tukey, J. W. (1958). Bias and confidence in not quite large samples. *Annals of Mathematical Statistics*, *29*, 614.

Walter, T., & Enge, P. (1995). Weighted RAIM for precision approach. *Proceedings of the 8th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1995)*, 1995–2004.

Wang, E., Jia, C., Tong, G., Qu, P., Lan, X., & Pang, T. (2018). Fault detection and isolation in GPS receiver autonomous integrity monitoring based on chaos particle swarm optimization-particle filter algorithm. *Advances in Space Research*, *61*(5), 1260–1272.

Wang, J., Xu, C., Shi, M., & Lu, Z. (2022). Protection level for precise point positioning based on Gaussian mixture model. *China Satellite Navigation Conference*, 45–55.

Wang, R., Xiong, Z., Liu, J., Xu, J., & Shi, L. (2016). Chi-square and SPRT combined fault detection for multisensor navigation. *IEEE Transactions on Aerospace and Electronic Systems*, *52*(3), 1352–1365.

Wen, W., Pfeifer, T., Bai, X., & Hsu, L. T. (2021). Factor graph optimization for GNSS/INS integration: A comparison with the extended Kalman filter. *NAVIGATION*, *68*(2), 315–331.

Yan, P., Wen, W., Huang, F., & Hsu, L.-T. (2024). A fault detection algorithm for LiDAR/IMU integrated localization systems with non-Gaussian noises. *To appear in Proceedings of the 2024 International Technical Meeting of The Institute of Navigation*. https://www.ion.org/itm/abstracts.cfm?paperID=12879

Yan, P., Zhong, Y., & Hsu, L.-T. (2024). Bounding the heavy-tailed DGNSS error by leveraging membership weights analysis of Gaussian mixture model. *To appear in Proceedings of the ION 2024 Pacific PNT Meeting*. https://www.ion.org/pnt/abstracts.cfm?paperID=13061

Yun, Y., Yun, H., Kim, D., & Kee, C. (2008). A Gaussian sum filter approach for DGNSS integrity monitoring. *The Journal of Navigation*, *61*(4), 687–703.

Zandbergen, R., Dinwiddy, S., Hahn, J., Breeuwer, E., & Blonski, D. (2004). Galileo orbit selection. *Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2004)*, 616–623.

Zhao, L., Zhang, J., Li, L., Yang, F., & Liu, X. (2020). Position-domain non-Gaussian error overbounding for ARAIM. *Remote Sensing*, *12*(12), 1992.

APPENDIX A. LINEARIZATION FOR SPP AND DGNSS POSITIONING SYSTEMS

1. Single Point Positioning (SPP)

The right-hand side (RHS) of the pseudorange measurement model in Equation (1) can be linearized by taking the first-order Taylor expansion at $\mathbf{x}_0 = [x_0, y_0, z_0, \delta_0]^T$ as follows:

$$\begin{aligned}\varrho_i = &\varrho_{i,0} - a_{i,1}(x - x_0) - a_{i,2}(y - y_0) - a_{i,3}(z - z_0) \\ &+ c(\delta_r - \delta_0) + \eta_i \,,\end{aligned} \tag{39}$$

where

$$\varrho_{i,0} = \sqrt{(p_x^i - x_0)^2 + (p_y^i - y_0)^2 + (p_z^i - z_0)^2} + c\delta_0 \tag{40a}$$

$$a_{i,1} = \frac{p_x^i - x_0}{\sqrt{(p_x^i - x_0)^2 + (p_y^i - y_0)^2 + (p_z^i - z_0)^2}} \tag{40b}$$

$$a_{i,2} = \frac{p_y^i - y_0}{\sqrt{(p_x^i - x_0)^2 + (p_y^i - y_0)^2 + (p_z^i - z_0)^2}} \tag{40c}$$

$$a_{i,3} = \frac{p_z^i - z_0}{\sqrt{(p_x^i - x_0)^2 + (p_y^i - y_0)^2 + (p_z^i - z_0)^2}} \,. \tag{40d}$$

The matrix form of the linearized pseudorange measurement model with n measurements can be written as

$$\mathbf{y} = \mathbf{H}\Delta\mathbf{x} + \boldsymbol{\varepsilon} \,, \tag{41}$$

where

$$\mathbf{y} = \begin{bmatrix} \varrho_{1,0} - \varrho_1 \\ \vdots \\ \varrho_{n,0} - \varrho_n \end{bmatrix}, \mathbf{H} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & 1 \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}, \Delta\mathbf{x} = \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \\ -c(\delta_r - \delta_0) \end{bmatrix} . \tag{42}$$

For single-frequency receivers (such as L1 or L5), the measurement error η_i mainly includes the satellite clock and ephemeris error $\epsilon_{eph\&clk,i}$, ionospheric delay $\epsilon_{iono,i}$, tropospheric delay $\epsilon_{trop,i}$ and code noise and multipath error $\epsilon_{cnmp,i}$:

$$\eta_i = \epsilon_{eph\&clk,i} + \epsilon_{iono,i} + \epsilon_{trop,i} + \epsilon_{cnmp,i} \,. \tag{43}$$

For double frequency receivers (e.g., L1-L5), the ionospheric delay can be eliminated by differencing pseudorange measurements on both frequencies as follows (Blanch et al., 2023; Misra & Enge, 2006),:

$$\varrho_{iono\text{-}free} = \frac{\gamma\varrho_{L1} - \varrho_{L5}}{\gamma - 1} \tag{44a}$$

$$\gamma = \left(\frac{f_{L1}}{f_{L5}}\right)^2 , \tag{44b}$$

where $\varrho_{iono\text{-}free}$ represents the ionospheric-free pseudorange, f_{L1} and f_{L5} represents the L1 and L5 frequencies, respectively. Then, the measurement error can be represented by (Blanch et al., 2023)

$$\eta_i = \epsilon_{eph\&clk,i}^{L1\text{-}L5} + \epsilon_{trop,i} + \frac{\gamma}{\gamma - 1}\epsilon_{cnmp,i}^{L1} - \frac{1}{\gamma - 1}\epsilon_{cnmp,i}^{L5} \,, \tag{45}$$

where $\epsilon_{eph\&clk,i}^{L1\text{-}L5}$ is the satellite clock and ephemeris error for the L1-L5 ionospheric free combination. With the definition in Equations (44) and (45), the linearized pseudorange measurement model for the dual frequency SPP takes a similar form as in Equations (41) and (42).

2. Differential Global Navigation Satellite System (DGNSS) Positioning

Equation (1) can be rewritten to represent the relationship between the receiver r_1 and the satellite s as follows:

$$\varrho^s_{r_1} = \|\mathbf{p}^s - \mathbf{x}_{r_1}\| + c\delta_{r_1} + \epsilon^s_{eph\&clk,r_1} + \epsilon^s_{iono,r_1} + \epsilon^s_{trop,r_1} + \epsilon^s_{cnmp,r_1}\,. \tag{46}$$

By taking the difference between the pseudoranges of receivers r_1 and r_2, the single-differenced pseudorange $\nabla\varrho^s_{r1,r2}$ is given by

$$\nabla\varrho^s_{r_1,r_2} = (\|\mathbf{p}^s - \mathbf{x}_{r_1}\| - \|\mathbf{p}^s - \mathbf{x}_{r_2}\|) + c\,(\delta_{r_1} - \delta_{r_2}) + \epsilon^s_{cnmp,r_1} - \epsilon^s_{cnmp,r_2}\,. \tag{47}$$

The common errors due to the satellite s are eliminated in the single difference. Similarly, for the measurements from satellite i,

$$\nabla\varrho^i_{r_1,r_2} = \left(\left\|\mathbf{p}^i - \mathbf{x}_{r_1}\right\| - \left\|\mathbf{p}^i - \mathbf{x}_{r_2}\right\|\right) + c\,(\delta_{r_1} - \delta_{r_2}) + \epsilon^i_{cnmp,r_1} - \epsilon^i_{cnmp,r_2}\,. \tag{48}$$

Then the double-differenced pseudorange is given by

$$\nabla\Delta\varrho^{s,i}_{r_1,r_2} = (\|\mathbf{p}^s - \mathbf{x}_{r_1}\| - \|\mathbf{p}^s - \mathbf{x}_{r_2}\|) - \left(\left\|\mathbf{p}^i - \mathbf{x}_{r_1}\right\| - \left\|\mathbf{p}^i - \mathbf{x}_{r_2}\right\|\right) + \epsilon^{s,i}_{r_1,r_2} \tag{49a}$$

$$\epsilon^{s,i}_{r_1,r_2} = \epsilon^s_{cnmp,r_1} - \epsilon^s_{cnmp,r_2} - \epsilon^i_{cnmp,r_1} + \epsilon^i_{cnmp,r_2}\,, \tag{49b}$$

where $\epsilon^{s,i}_{r_1,r_2}$ is the remaining error after double differencing and is called the DGNSS error.

Assuming that $\mathbf{x}_{r_2}$ is known and $\mathbf{x} = \mathbf{x}_{r_1} = [x, y, z]^T$ is the position to be estimated, the right-hand side (RHS) of (49a) can be linearized by taking the first-order Taylor expansion at $\mathbf{x}_0 = [x_0, y_0, z_0]^T$ as follows:

$$\nabla\Delta\varrho^{s,i}_{r_1,r_2} = \nabla\Delta\hat{\varrho}^{s,i}_{r_1,r_2} - e_{i,1}\,(x - x_0) - e_{i,2}\,(y - y_0) - e_{i,3}\,(z - z_0) + \epsilon^{s,i}_{r_1,r_2}\,, \tag{50}$$

where

$$\nabla\Delta\hat{\varrho}^{s,i}_{r_1,r_2} = (\|\mathbf{p}^s - \mathbf{x}_0\| - \|\mathbf{p}^s - \mathbf{x}_{r_2}\|) - \left(\left\|\mathbf{p}^i - \mathbf{x}_0\right\| - \left\|\mathbf{p}^i - \mathbf{x}_{r_2}\right\|\right) \tag{51a}$$

$$e_{i,1} = \frac{p^s_x - x_0}{\|\mathbf{p}^s - \mathbf{x}_0\|} - \frac{p^i_x - x_0}{\|\mathbf{p}^i - \mathbf{x}_0\|} \tag{51b}$$

$$e_{i,2} = \frac{p^s_y - y_0}{\|\mathbf{p}^s - \mathbf{x}_0\|} - \frac{p^i_y - y_0}{\|\mathbf{p}^i - \mathbf{x}_0\|} \tag{51c}$$

$$e_{i,3} = \frac{p^s_z - z_0}{\|\mathbf{p}^s - \mathbf{x}_0\|} - \frac{p^i_z - z_0}{\|\mathbf{p}^i - \mathbf{x}_0\|}\,. \tag{51d}$$

The matrix form of the linearized double-differenced pseudorange measurement model with n measurements can be written as

$$\mathbf{y} = \mathbf{H}\Delta\mathbf{x} + \boldsymbol{\varepsilon}\,, \tag{52}$$

where

$$\mathbf{y} = \begin{bmatrix} \nabla\Delta\hat{\varrho}^{s,1}_{r_1,r_2} - \nabla\Delta\varrho^{s,1}_{r_1,r_2} \\ \vdots \\ \nabla\Delta\hat{\varrho}^{s,n}_{r_1,r_2} - \nabla\Delta\varrho^{s,n}_{r_1,r_2} \end{bmatrix}, \mathbf{H} = \begin{bmatrix} e_{1,1} & e_{1,2} & e_{1,3} \\ \vdots & \vdots & \vdots \\ e_{n,1} & e_{n,2} & e_{n,3} \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \epsilon^{s,1}_{r_1,r_2} \\ \vdots \\ \epsilon^{s,n}_{r_1,r_2} \end{bmatrix}, \Delta\mathbf{x} = \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix}. \tag{53}$$

APPENDIX B. DISTRIBUTION OF JACKKNIFE RESIDUAL UNDER GAUSSIAN NOISES

The Gauss-Markov conditions concern the set of noises in the linear system $\mathbf{y} = \mathbf{H}\Delta\mathbf{x} + \boldsymbol{\varepsilon}$ as follows:

1. Zero mean: $\mathrm{E}[\varepsilon_i] = 0\ \forall i$;
2. Homoscedastic: $\mathrm{Var}[\varepsilon_i] = \sigma^2 < \infty\ \forall i$;
3. Uncorrelated: $\mathrm{Cov}[\varepsilon_i, \varepsilon_j] = 0\ \forall i \neq j$.

Under Gauss-Markov conditions, the ordinary least squares (OLS) estimator is the best linear unbiased estimator (BLUE).

A further generalization of the Gauss-Markov conditions to heteroscedastic and correlated errors was developed (Aitken, 1936), and its application to the weighted least squares (WLS) estimator can be stated as follows:

"WLS is the BLUE if the weight matrix is equal to the inverse of the variance-covariance matrix of the noises."

Based on the generalized Gauss-Markov conditions, the subsolution in (9a) has the following properties:

$$\mathrm{E}[\Delta\hat{\mathbf{x}}^{(i)}]=0 \tag{54a}$$

$$\mathrm{Var}[\Delta\hat{\mathbf{x}}^{(i)}]=\mathbf{S}^{(i)}\mathbf{W}^{-1}{\mathbf{S}^{(i)}}^{T}\,. \tag{54b}$$

Therefore, the expectation and variance of the Jackknife residual in Equation (14) are given by

$$\mathrm{E}[t_i]=0 \tag{55a}$$

$$\begin{aligned}\mathrm{Var}[t_i]&=\mathbf{g}_i\,\mathrm{Var}\left[\Delta\mathbf{x}^{(i)}-\Delta\hat{\mathbf{x}}^{(i)}\right]\mathbf{g}_i^T+\sigma_i^2\\&=\mathbf{g}_i\mathbf{S}^{(i)}\mathbf{W}^{-1}{\mathbf{S}^{(i)}}^{T}\mathbf{g}_i^T+\sigma_i^2\,.\end{aligned} \tag{55b}$$

As shown in Equation (29), the Jackknife residual can be rewritten as

$$t_i=\tilde{\mathbf{p}}_i\boldsymbol{\varepsilon}\,, \tag{56}$$

which is a linear combination of measurement noises. If ε_i has a zero-mean Gaussian distribution defined in Equation (15), t_i will have a Gaussian distribution

$$t_i\sim\mathcal{N}\left(0,\tilde{\mathbf{p}}_i\mathbf{W}^{-1}\tilde{\mathbf{p}}_i^T\right)\,. \tag{57}$$

Since a Gaussian distribution is uniquely defined by its mean and variance, the following equation will hold:

$$\mathbf{g}_i\mathbf{S}^{(i)}\mathbf{W}^{-1}{\mathbf{S}^{(i)}}^{T}\mathbf{g}_i^T+\sigma_i^2=\tilde{\mathbf{p}}_i\mathbf{W}^{-1}\tilde{\mathbf{p}}_i^T\,. \tag{58}$$

Therefore,

$$t_i\sim\mathcal{N}\left(0,\mathbf{g}_i\mathbf{S}^{(i)}\mathbf{W}^{-1}{\mathbf{S}^{(i)}}^{T}\mathbf{g}_i^T+\sigma_i^2\right)\,. \tag{59}$$

APPENDIX C. BONFERRONI CORRECTION

The hypotheses with Bonferroni correction (Bonferroni, 1936) in Equation (18) have the following relationship with the original hypotheses in Equation (17):

$$\begin{aligned}H_0&=\bigcap_{i=1}^{n}H_0^{(i)}\\H_1&=\bigcup_{i=1}^{n}H_1^{(i)}\,.\end{aligned} \tag{60}$$

Assume that the probability of type I error of the corrected hypothesis test is α^*. Then,

$$\begin{aligned}1-\alpha^*&=P(\text{All tests accept}|H_0)\\&=1-P(\text{At least one test is rejected}|H_0)\\&\geq 1-\sum_{i=1}^{n}P(\text{Origin test }i\text{ is rejected}|H_0)\\&=1-\sum_{i=1}^{n}P(\text{Origin test }i\text{ is rejected}|H_0^{(i)})\\&=1-n\alpha\,.\end{aligned} \tag{61}$$

In addition,

$$\begin{aligned}\alpha^*&=P(\text{At least one test is rejected}|H_0)\\&\geq P(\text{Origin test }i\text{ is rejected}|H_0^{(i)})\\&=\alpha\,.\end{aligned} \tag{62}$$

Therefore,

$$\alpha \leq \alpha^* \leq n\alpha\,. \tag{63}$$

To keep the type I error α^* not exceeding τ (e.g., 0.05),

$$n\alpha = \tau\,. \tag{64}$$

Thus, the type I error of the individual test would be $\alpha = \frac{\tau}{n}$.

APPENDIX D. DERIVATION IN THE EXTENDED JACKKNIFE RESIDUAL

Substitute Equation (8) and Equation (22) into Equation (24)

$$\begin{aligned}\tilde{\mathbf{P}} &= \begin{bmatrix}\mathbf{H}^{(i,*)}_{1:i-1,:}\\ \mathbf{g}_i\\ \mathbf{H}^{(i,*)}_{i:n-1,:}\end{bmatrix}\begin{bmatrix}\mathbf{S}^{(i,*)}_{:,1:i-1} & 0 & \mathbf{S}^{(i,*)}_{:,i:n-1}\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{H}^{(i,*)}_{1:i-1,:}\mathbf{S}^{(i,*)}_{:,1:i-1} & 0 & \mathbf{H}^{(i,*)}_{1:i-1,:}\mathbf{S}^{(i,*)}_{:,i:n-1}\\ \mathbf{g}_i\mathbf{S}^{(i,*)}_{:,1:i-1} & 0 & \mathbf{g}_i\mathbf{S}^{(i,*)}_{:,i:n-1}\\ \mathbf{H}^{(i,*)}_{i:n-1,:}\mathbf{S}^{(i,*)}_{:,1:i-1} & 0 & \mathbf{H}^{(i,*)}_{i:n-1,:}\mathbf{S}^{(i,*)}_{:,i:n-1}\end{bmatrix}.\end{aligned} \tag{65}$$

Then

$$\left(\mathbf{I}-\tilde{\mathbf{P}}\right)\tilde{\mathbf{H}}^{(i)} = \begin{bmatrix}\mathbf{I}-\mathbf{H}^{(i,*)}_{1:i-1,:}\mathbf{S}^{(i,*)}_{:,1:i-1} & 0 & -\mathbf{H}^{(i,*)}_{1:i-1,:}\mathbf{S}^{(i,*)}_{:,i:n-1}\\ -\mathbf{g}_i\mathbf{S}^{(i,*)}_{:,1:i-1} & 1 & -\mathbf{g}_i\mathbf{S}^{(i,*)}_{:,i:n-1}\\ -\mathbf{H}^{(i,*)}_{i:n-1,:}\mathbf{S}^{(i,*)}_{:,1:i-1} & 0 & \mathbf{I}-\mathbf{H}^{(i,*)}_{i:n-1,:}\mathbf{S}^{(i,*)}_{:,i:n-1}\end{bmatrix}\begin{bmatrix}\mathbf{H}^{(i,*)}_{1:i-1,:}\\ \mathbf{g}_i\\ \mathbf{H}^{(i,*)}_{i:n-1,:}\end{bmatrix}. \tag{66}$$

Based on the property of block matrix multiplication, it is equivalent to write Equation (66) as

$$\left(\mathbf{I}-\tilde{\mathbf{P}}\right)\tilde{\mathbf{H}}^{(i)} = \begin{bmatrix}\mathbf{I}-\mathbf{H}^{(i,*)}_{1:i-1,:}\mathbf{S}^{(i,*)}_{:,1:i-1} & -\mathbf{H}^{(i,*)}_{1:i-1,:}\mathbf{S}^{(i,*)}_{:,i:n-1} & 0\\ -\mathbf{g}_i\mathbf{S}^{(i,*)}_{:,1:i-1} & -\mathbf{g}_i\mathbf{S}^{(i,*)}_{:,i:n-1} & 1\\ -\mathbf{H}^{(i,*)}_{i:n-1,:}\mathbf{S}^{(i,*)}_{:,1:i-1} & \mathbf{I}-\mathbf{H}^{(i,*)}_{i:n-1,:}\mathbf{S}^{(i,*)}_{:,i:n-1} & 0\end{bmatrix}\begin{bmatrix}\mathbf{H}^{(i,*)}_{1:i-1,:}\\ \mathbf{H}^{(i,*)}_{i:n-1,:}\\ \mathbf{g}_i\end{bmatrix}. \tag{67}$$

By introducing a set of permutation matrix $\mathbf{P}_{i,i+1}, \mathbf{P}_{i+1,i+2}, \cdots, \mathbf{P}_{n-1,n}$ ($\mathbf{P}_{i,j}$ represents the permutation matrix to exchange ith and jth row), the ith row of the first matrix on the right-hand-side of Equation (67) can be moved to the nth row, as shown below,

$$\begin{aligned}\left(\mathbf{I}-\tilde{\mathbf{P}}\right)\tilde{\mathbf{H}}^{(i)} &= \mathbf{P}^{-1}_{i,i+1}\mathbf{P}^{-1}_{i+1,i+2}\cdots\mathbf{P}^{-1}_{n-1,n}\begin{bmatrix}\mathbf{I}-\mathbf{H}^{(i,*)}_{1:i-1,:}\mathbf{S}^{(i,*)}_{:,1:i-1} & -\mathbf{H}^{(i,*)}_{1:i-1,:}\mathbf{S}^{(i,*)}_{:,i:n-1} & 0\\ -\mathbf{H}^{(i,*)}_{i:n-1,:}\mathbf{S}^{(i,*)}_{:,1:i-1} & \mathbf{I}-\mathbf{H}^{(i,*)}_{i:n-1,:}\mathbf{S}^{(i,*)}_{:,i:n-1} & 0\\ -\mathbf{g}_i\mathbf{S}^{(i,*)}_{:,1:i-1} & -\mathbf{g}_i\mathbf{S}^{(i,*)}_{:,i:n-1} & 1\end{bmatrix}\begin{bmatrix}\mathbf{H}^{(i,*)}_{1:i-1,:}\\ \mathbf{H}^{(i,*)}_{i:n-1,:}\\ \mathbf{g}_i\end{bmatrix}\\ &= \mathbf{P}^{-1}_{i,i+1}\mathbf{P}^{-1}_{i+1,i+2}\cdots\mathbf{P}^{-1}_{n-1,n}\begin{bmatrix}\mathbf{I}-\mathbf{H}^{(i,*)}\mathbf{S}^{(i,*)} & 0\\ -\mathbf{g}_i\mathbf{S}^{(i,*)} & 1\end{bmatrix}\begin{bmatrix}\mathbf{H}^{(i,*)}\\ \mathbf{g}_i\end{bmatrix}\\ &= \mathbf{P}^{-1}_{i,i+1}\mathbf{P}^{-1}_{i+1,i+2}\cdots\mathbf{P}^{-1}_{n-1,n}\begin{bmatrix}\mathbf{H}^{(i,*)}-\mathbf{H}^{(i,*)}\mathbf{S}^{(i,*)}\mathbf{H}^{(i,*)}\\ -\mathbf{g}_i\mathbf{S}^{(i,*)}\mathbf{H}^{(i,*)}+\mathbf{g}_i\end{bmatrix}.\end{aligned} \tag{68}$$

According to the definition of $\mathbf{S}^{(i,*)}$ in Equation (7), it is easily to found that

$$\mathbf{S}^{(i,*)}\mathbf{H}^{(i,*)} = \mathbf{I}\,. \tag{69}$$

Therefore,

$$\left(\mathbf{I}-\tilde{\mathbf{P}}\right)\tilde{\mathbf{H}}^{(i)} = 0\,. \tag{70}$$

APPENDIX E. WEIGHTED SUM OF INDEPENDENT RANDOM VARIABLES

Let $Y_j = \tilde{p}_{i,j}\varepsilon_j$ be an independent random variable. Equation (30) can be written as

$$t_i = \sum_{j=1}^{n} Y_j \,. \tag{71}$$

Since the PDF of the sum of independent variables is the convolution of their PDFs, the PDF of t_i can be written as

$$f_{t_i}(x) = f_{Y_1}(x) * f_{Y_2}(x) * \ldots * f_{Y_n}(x) \,, \tag{72}$$

where $*$ denotes the convolution operation. The CDF of Y_j can be written as

$$\begin{aligned} F_{Y_j}(y) &= P\left(Y_j < y\right) = P\left(\varepsilon_j < \frac{y}{|\tilde{p}_{i,j}|}\right) \\ &= \int_{-\infty}^{y/|\tilde{p}_{i,j}|} f_{\varepsilon_j}(x)dx \,. \end{aligned} \tag{73}$$

Let $t = |\tilde{p}_{i,j}|\, x$, then

$$F_{Y_j}(y) = \frac{1}{|\tilde{p}_{i,j}|}\int_{-\infty}^{y} f_{\varepsilon_j}\left(\frac{t}{|\tilde{p}_{i,j}|}\right)dt \,. \tag{74}$$

By taking the first derivative with respect to t, the PDF of Y_j can be obtained as follows:

$$f_{Y_j}(x) = \frac{1}{|\tilde{p}_{i,j}|} f_{\varepsilon_j}\left(\frac{x}{|\tilde{p}_{i,j}|}\right) \,. \tag{75}$$

Therefore, Equation (72) can be written as,

$$f_{t_i}(x) = \prod_{j=1}^{n} |\tilde{p}_{i,j}|^{-1} f_{\varepsilon_1}\left(\frac{x}{|\tilde{p}_{i,1}|}\right) * f_{\varepsilon_2}\left(\frac{x}{|\tilde{p}_{i,2}|}\right) * \ldots * f_{\varepsilon_n}\left(\frac{x}{|\tilde{p}_{i,n}|}\right) \,. \tag{76}$$

APPENDIX F. PRINCIPAL GAUSSIAN OVERBOUND

The construction of Principal Gaussian Overbound (PGO) mainly consists of the following three steps:

(1) Error fitting with zero-mean BGMM

Let the fitted zero-mean bimodal Gaussian mixture model (BGMM) for the error distribution be

$$f(x) = p_1 f_{\mathcal{N}}\left(x; 0, \delta_1^2\right) + \left(1 - p_1\right) f_{\mathcal{N}}\left(x; 0, \delta_2^2\right) \,, \tag{77}$$

where $f_{\mathcal{N}}\left(x; 0, \delta_1^2\right)$ and $f_{\mathcal{N}}\left(x; 0, \delta_2^2\right)$ are the PDF of the first and the second Gaussian component, δ_1 and δ_2 the corresponding standard deviations, and p_1 and $1 - p_1$ are the mixing weight of the two Gaussian components, respectively.

(2) Dominance partition

The membership weights s_1 and s_2 of Equation (77) can be written as

$$s_1(x) = \frac{p_1 f_{\mathcal{N}}\left(x; 0, \delta_1^2\right)}{p_1 f_{\mathcal{N}}\left(x; 0, \delta_1^2\right) + \left(1 - p_1\right) f_{\mathcal{N}}\left(x; 0, \delta_2^2\right)} \tag{78a}$$

$$s_2(x) = \frac{\left(1 - p_1\right) f_{\mathcal{N}}\left(x; 0, \delta_2^2\right)}{p_1 f_{\mathcal{N}}\left(x; 0, \delta_1^2\right) + \left(1 - p_1\right) f_{\mathcal{N}}\left(x; 0, \delta_2^2\right)} \,. \tag{78b}$$

The membership weights indicate the posterior probability of a data point being generated from each component. It is proved in Yan, Zhong, and Hsu (2024) that $s_1(x)$ is a symmetric concave function and $s_2(x)$ is a symmetric convex function. Figure

7(a) shows an example plot of $s_1(x)$ and $s_2(x)$ for a zero-mean BGMM with $p_1 = 0.9$, $\delta_1 = 0.5$, and $\delta_2 = 1$. It is shown that $s_1(x)$ exhibits large values when x is situated at the core region of the BGMM. As x goes far away from the center, the value diminishes rapidly. Conversely, $s_2(x)$ exhibits an opposite trend, with its highest value occurring in the tail region. These trends illustrate the dominance of each Gaussian component in BGMM. Therefore, a partitioning strategy is constructed by solving the equations as follows:

$$s_2\left(x_{lp}\right) = \alpha \cdot \lim_{t\to\infty} s_2(t) \; \forall x_{lp} < 0 \tag{79a}$$

$$s_2\left(x_{rp}\right) = \alpha \cdot \lim_{t\to\infty} s_2(t) \; \forall x_{rp} > 0\,, \tag{79b}$$

where x_{lp} and x_{rp} are the core-tail transition points, $x \in [x_{lp}, x_{rp}]$ defines the core region, $x \in [-\infty, x_{lp}] \cup [x_{rp}, \infty]$ defines the tail region, and $\alpha \in (0.5, 1)$ is the partition parameter describing the degree of dominance of the 2nd Gaussian component in the tail region.

With this partition, data points located in the tail region demonstrate a larger membership weight in the 2nd Gaussian component compared to x_{lp} or x_{rp}. Notably, this value increases monotonically as the data point moves farther away from the center, indicating the dominance of the 2nd Gaussian component in the tail region. However, in the core region, the situation becomes complicated. The intersection points of $s_1(x)$ and $s_2(x)$ can be computed as follows:

$$x^{\mathrm{L}}_{\mathrm{intersect}} = -\sqrt{\frac{2\delta_1^2\delta_2^2}{\delta_2^2 - \delta_1^2} \ln \frac{p_1\delta_2}{(1-p_1)\,\delta_1}} \tag{80a}$$

$$x^{\mathrm{R}}_{\mathrm{intersect}} = \sqrt{\frac{2\delta_1^2\delta_2^2}{\delta_2^2 - \delta_1^2} \ln \frac{p_1\delta_2}{(1-p_1)\,\delta_1}}\,. \tag{80b}$$

The interval $\left[x^{\mathrm{L}}_{\mathrm{intersect}},\, x^{\mathrm{R}}_{\mathrm{intersect}}\right]$ can be interpreted as the region that the 1st component has more contribution than the 2nd component. When p_1 goes larger, $x^{\mathrm{R}}_{\mathrm{intersect}}$ will increase; thereby, the length of the interval $\left[x^{\mathrm{L}}_{\mathrm{intersect}},\, x^{\mathrm{R}}_{\mathrm{intersect}}\right]$ will be extended. In other words, the probability of the 1st component's dominance in the core region becomes larger. Then, it is reasonable to assume that the core region is dominated by the 1st component. This assumption is rational in the context of heavy-tailed distribution modeling, especially for pseudorange error modeling.

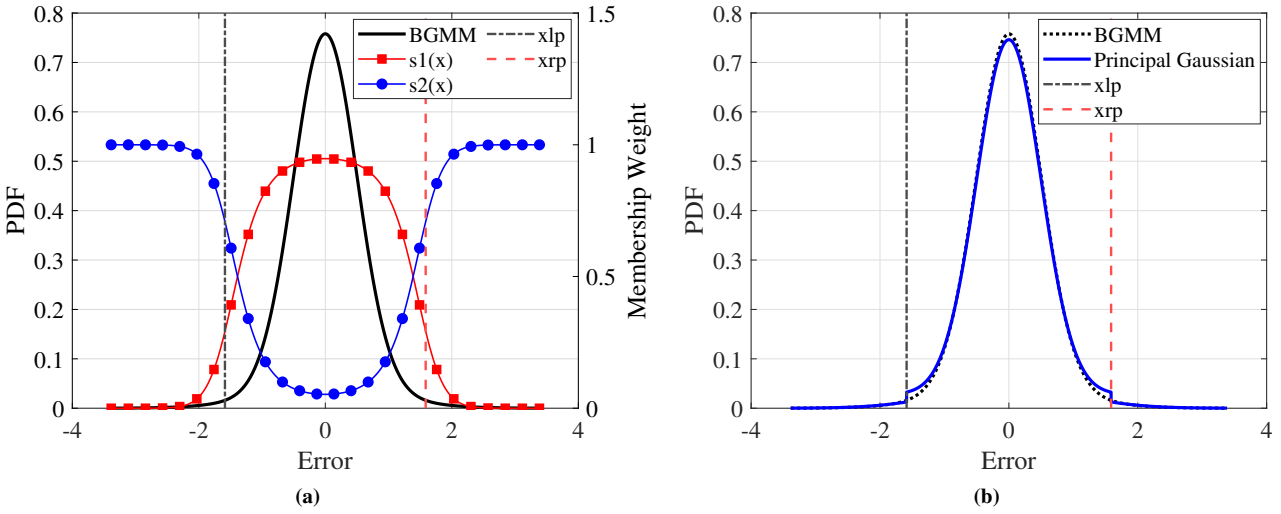


Figure 7: (a) Membership weights of a zero-mean bimodal Gaussian mixture model with $p_1 = 0.9$, $\delta_1 = 0.5$, and $\delta_2 = 1$. (b) PDF of the Principal Gaussian overbound and the Gaussian overbound for the zero-mean bimodal Gaussian mixture model. x_{lp} refers to $x^{\mathrm{L}}_{\mathrm{intersect}}$, and x_{rp} refers to $x^{\mathrm{R}}_{\mathrm{intersect}}$.

(3) Overbounding each region

Within each region, a CDF overbound is constructed based on the inflation and shifting of the dominant Gaussian component

with the constraints of unimodality. The PDF of the finalized overbound distribution is given by (Yan, Zhong, & Hsu, 2024)

$$f_{PGO}(x) = \begin{cases} (1+k)\,(1-p_1)\,f_{\mathcal{N}}\left(x;0,\delta_2^2\right) & |x| > x_{rp} \\ p_1 f_{\mathcal{N}}\left(x;0,\delta_1^2\right) + c & |x| \le x_{rp} \end{cases}, \tag{81}$$

where

$$k = \frac{p_1 G\left(x_{lp};0,\delta_1\right)}{(1-p_1)\,G\left(x_{lp};0,\delta_2\right)}, c = \frac{(1-p_1)\left(G\left(x_{lp};0,\delta_2\right)-0.5\right)}{x_{lp}}. \tag{82}$$

It is not difficult to find that k and c are solely determined by the partition parameter α. The PDF of the PGO for the zero-mean BGMM ($p_1 = 0.9$, $\delta_1 = 0.5$, $\delta_2 = 1$) is plotted in Figure 7(b), which is shown as a piece-wised function. The plot reveals that the PDF of the PGO is closely aligned with those of BGMM in both the tail and core regions.

APPENDIX G. GAUSSIAN ERROR MODELS FOR WORLDWIDE SIMULATIONS

This paper uses the setting in Blanch et al. (2010) to set the Gaussian error models for the dual-frequency SPP. Specifically, the error bound for the satellite clock and ephemeris error for the L1-L5 ionospheric free combination is set as

$$\sigma_{eph\&clk,i}^{L1\text{-}L5} = 0.5\,\mathrm{m}. \tag{83}$$

The tropospheric error bound is assumed to be

$$\sigma_{trop,i} = (0.12\,\mathrm{m})\Big(\frac{1.001}{\sqrt{0.002001+\sin^2\theta_i}}\Big). \tag{84}$$

The code noise and multipath error bound for L1 frequency is given by

$$\left(\sigma_{cnmp,i}^{L1}\right)^2 = \left(\sigma_{noise,i}^{L1}\right)^2 + \left(\sigma_{multipath,i}^{L1}\right)^2, \tag{85}$$

where the code noise error bound is given by (Oehler et al., 2004)

$$\sigma_{noise,i}^{L1} = 0.04\,\mathrm{m} - (0.02\,\mathrm{m})\frac{\theta_i - 5^\circ}{85^\circ}, \tag{86}$$

and the multipath error bound is defined for AAD-A aircraft, as shown in Rife, Pullen, Pervan, and Enge (2004)

$$\sigma_{multipath,i}^{L1} = 0.13\,\mathrm{m} + (0.53\,\mathrm{m})\exp\Big(-\frac{\theta_i}{10^\circ}\Big). \tag{87}$$

The code noise and multipath error bound for L5 frequency is set to be the same as that of L1 frequency,

$$\sigma_{cnmp,i}^{L1} = \sigma_{cnmp,i}^{L5}. \tag{88}$$

APPENDIX H. RESULTS OF OVERBOUNDING CORS DGNSS ERRORS

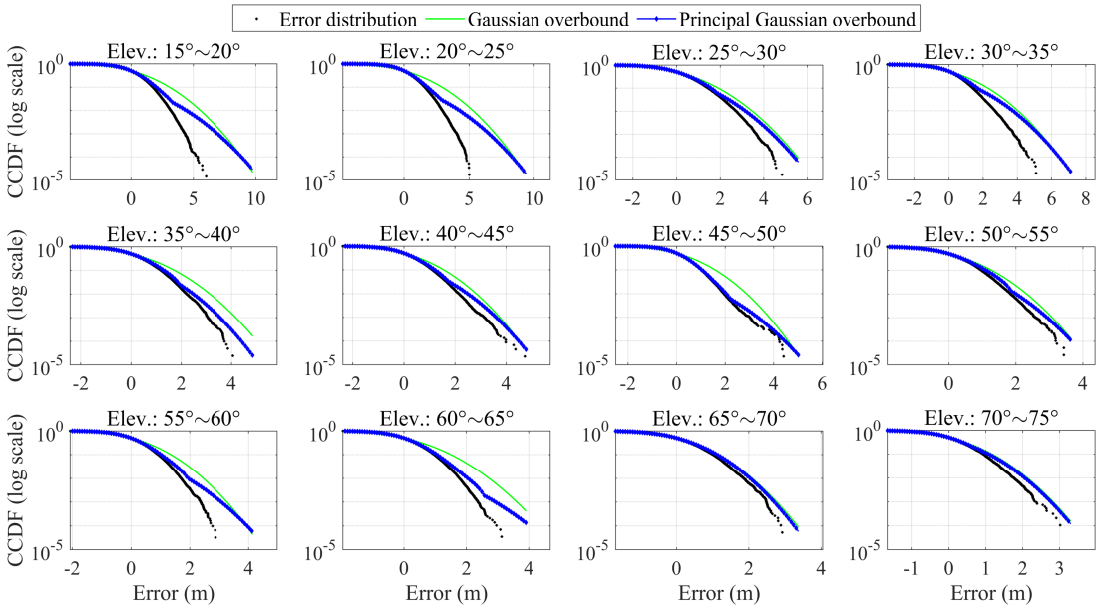


Figure 8: The complementary cumulative distribution function (CCDF) of the Gaussian overbound and Principal Gaussian overbound for DGNSS errors in each elevation angle bin. The CCDF is plotted in logarithm scale.

APPENDIX I. RESULTS OF OVERBOUNDING CORRECTED PSEUDORANGE ERRORS AT CHTI STATION

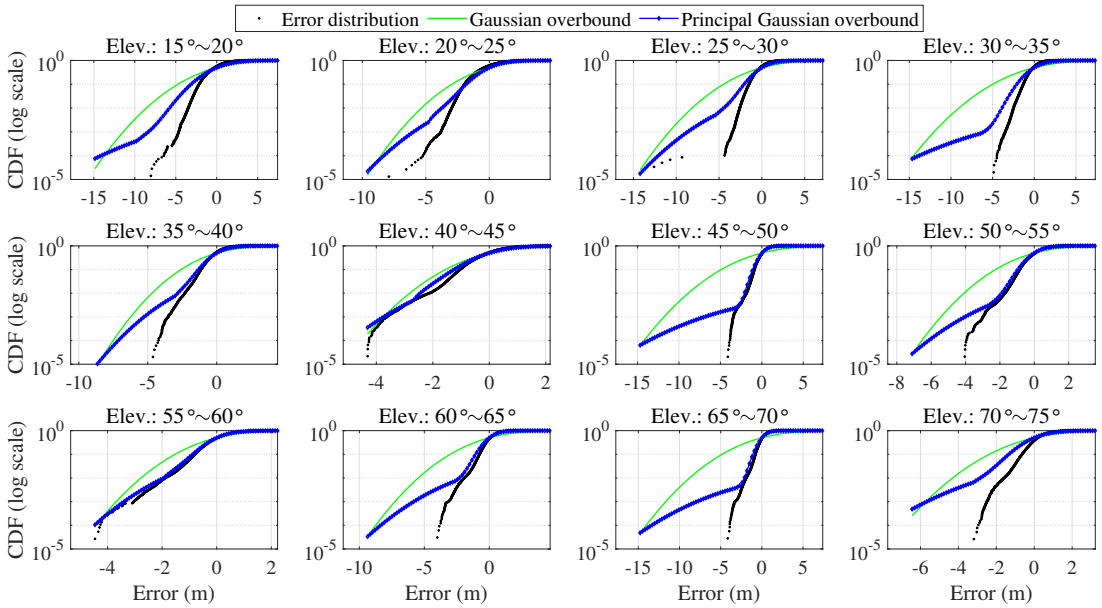


Figure 9: The CDF of the Gaussian overbound and Principal Gaussian overbound for corrected pseudorange errors in each elevation angle bin. The CDF is plotted in logarithm scale.

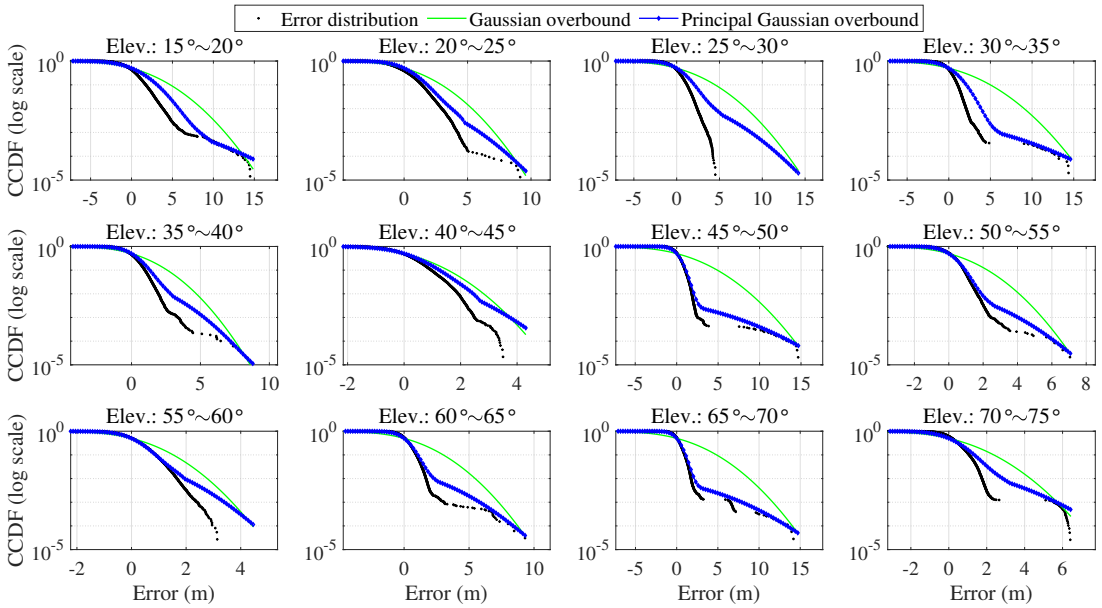


Figure 10: The complementary cumulative distribution function (CCDF) of the Gaussian overbound and Principal Gaussian overbound for corrected pseudorange errors in each elevation angle bin. The CCDF is plotted in logarithm scale.