Single Frequency GNSS Carrier Phase Cycle Slip Detection and Identification Using a Factor Graph Approach

Hakim Cherfi, Julien Lesouple, Joan Solà, Paul Thevenon

Single Frequency GNSS Carrier Phase Cycle Slip Detection and Identification Using a Factor Graph Approach

Hakim Cherfi, *Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, France.*
Julien Lesouple, *Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, France.*
Joan Solà, *Institut de Robòtica i Informàtica Industrial – CSIC, Barcelona, Catalonia.*
Paul Thevenon, *Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, France.*

BIOGRAPHY

Hakim Cherfi received his Eng. degree in robotics from UPSSITECH, Toulouse, France in 2022. He started his PhD in October 2022 at ENAC, Toulouse, within the SIGNAV team. His work focuses on precise state estimation of smartphones using GNSS and inertial sensors.

Julien Lesouple received the Eng. degree in Aeronautics Engineering from ISAE ENSICA, Toulouse, France in 2014 and defended in March 2019 his Ph.D degree in Signal Processing applied to GNSS at the cooperative laboratory TéSA. Since October 2021, he has been a lecturer researcher at ENAC within SIGNAV team. His research interests include statistical signal processing with applications to position, navigation and timing.

Joan Solà received his Eng. degree in Telecommunications by UPC-Barcelona in 1995, and Master and PhD in Automatic Control by the University of Toulouse in 2003 and 2007, hosted by LAAS-CNRS. He was a fellow scientist at SRI-International (California) in 2007. Since 2014 he is senior scientist with the Institut de Robòtica i Informàtica Industrial in Barcelona (IRI). Joan has devoted most of his research to state estimation, with contributions to monocular SLAM and high-dynamics motion. Joan has also experience in the industry in the fields of power electronics and deep-water submarine technology. He is currently leading activities at IRI regarding high-dynamic robotics involving estimation, control and robot design.

Paul Thevenon graduated as electronic engineer from Ecole Centrale de Lille in 2004 and obtained in 2007 a research master at ISAE in space telecommunications and a PhD degree in 2010 in the signal processing at ENAC. From 2010 to 2013, he was employed by CNES to supervise GNSS research activities and measurement campaigns. Since the July 2013, he is employed by ENAC as Assistant Professor. His current activities are GNSS signal quality monitoring and GNSS precise positioning algorithms.

ABSTRACT

Precise and robust GNSS positioning with low-cost receivers is a challenge, as these receivers will often be disturbed by biased observations due to a given environment, for example by code pseudoranges with multipath. One way to reach centimeter positioning accuracy is by using very precise observations, in particular by using GNSS carrier phase measurements. These carrier phase measurements can be very precise (under centimeter standard deviation) but they contain an unknown, the phase ambiguity. Throughout time, for a tracked satellite during multiple epochs, this phase ambiguity is supposed to remain constant. So by building an observation model containing observations from multiple epochs, we can take advantage of this information. This is what is usually done in a discrete-time Kalman filter: we estimate the ambiguity of each satellite, and we model this ambiguity to be constant in the state transition model. However, in practice, cycle slips happen. If we suppose that the ambiguity is constant, then the positioning algorithms solutions will be degraded when cycle slips happen.

This work uses factor graph optimization in order to detect and identify cycle slips on GNSS carrier phase observations from a single frequency receiver. This is done in a two-step process. It starts with a theoretical and very general approach to detect faults on measurements, as a binary decision. This is done with a theoretical distribution of a residual-based statistic (difference between observation models and measurements), and a comparison of the observed statistic with its theoretical distribution. In the case of detected fault, then comes step two of the process: by building multiple models and an optimization problem over a finite and discrete set of hypotheses, we are able to identify the observation where a fault was introduced. Following this theoretical approach to detect and identify faults, factor graph optimization is introduced. This is the strength of our approach: we are able to apply the procedure of detection and identification of faults for any factor graph, i.e., for any set of observations and states to estimate. Finally, the detection method is implemented in a GNSS simulator. We build a factor graph that estimates a receiver position from multiple carrier phase measurements. We apply the detection method to detect cycle slips on phase measurements. We show that, depending on the duration of the observation window, our detection method works for very small cycle slips, of one wavelength. We also compare Monte-Carlo simulations against theoretical probabilities of detection.

https://doi.org/10.33012/2024.19642

I. INTRODUCTION

Carrier phase GNSS observations are of upmost interest for precise positioning and navigation applications, since these observations are very precise, compared to code range measurements [Roberts, 2019]. However, these observations are ambiguous, meaning that for a given wavelength corresponding to a given carrier phase frequency, and for a given satellite, the observations are corrupted by an integer number of wavelengths [Wang et al., 2022]. Most precise positioning applications will assume that this integer number of wavelengths remains constant throughout time for a tracked satellite, and will use this information in order to estimate this number, leading to precise positioning. Unfortunately, this integer number can change over time and make the assumptions wrong, leading to a so called "cycle slip". A cycle slip occurs due to loss of lock of a receiver's phase tracking loops [Langley, 2014] that can be due to multiple reasons: obstruction of signal, a low C/N_0, an interference, or even a bug in the receiver software [Hofmann-Wellenhof et al., 2008]. Many applications exist that rely on constant ambiguities. Precise Point Positioning (PPP) assumes a constant transition model for the ambiguities in a Kalman filter, along with other transition models for the other state variables [Lotfy et al., 2022]. Other authors use this assumption to build temporal differences between carrier phase observations without ambiguity in their models [Kirkko-Jaakkola et al., 2009].

Regarding cycle slips, many authors have proposed techniques to detect and correct them. For detection purposes, most of them [Farooq et al., 2018] build a quantity and use it to make a decision. Hopefully, the test will provide the lowest probability of false alarm and the highest probability of detection. For example, in a Kalman filter, this is typically done by computing the norm of the innovation vector [Kerhascoet et al., 2017], [Hajiyev, 2016]. In the GNSS community, there are ways to improve this decision: some authors use carrier phase observations at multiple frequencies, and by combining observations, they are able to increase the wavelength, leading to larger cycle slips [Li and Melachroinos, 2019]. Others build larger observation systems containing fault-free measurements, in order to increase their detection capability [Wen and Hsu, 2022], and some use geometric constraints [Knowles and Gao, 2023]. The cycle slip detector proposed in this article relies on the factor graph framework. The factor graph framework has been used in many applications. For example, in GNSS positioning [Wen and Hsu, 2021], in GNSS odometry [Congram and Barfoot, 2022], for hybridization of sensors [Bai et al., 2022]. From the factor graph outcome, an hypothesis testing is proposed to detect the presence of a faulty measurement. This principle is then transposed to a cycle slip detector, using carrier phase observations. Then, if a fault is detected, a new set of factor graphs is generated to catch the faulty satellite. The motivation behind this theory is that most cycle slip detectors can be modelled by a factor graph. We will start by a very general theory of least-squares estimation, followed by statistics on residuals. We will propose a set of conditions on faults that we will be able to detect and exclude, followed by the actual tests to detect and identify them. Then, the factor graph framework will be introduced, and we will see the link with the previously explained theory. Finally, we will propose a factor graph made of GNSS carrier phase observations, and will see how our theory can detect cycle slips.

II. LINEAR SYSTEM, ESTIMATION AND RESIDUAL

The following sections are very general, as done in [Joerger et al., 2014], but will be the building blocks of the proposed factor graph detector. Readers aware of Receiver Autonomous Integrity Monitoring (RAIM) and integrity monitoring might be familiar with the following sections [Kirkko-Jaakkola et al., 2009].

1. Linear system estimation

We start with a full-rank linear observation system:

$$\boldsymbol{y}_R = \boldsymbol{H}_R \boldsymbol{x} + \boldsymbol{\varepsilon}_R \tag{1}$$

with $\boldsymbol{x} \in \mathbb{R}^n$ a hidden state vector to estimate, $\boldsymbol{y}_R \in \mathbb{R}^m$ a known measurement vector, $\boldsymbol{H}_R \in \mathcal{M}_{m,n}(\mathbb{R})$ with $m \geq n$ a known full-rank observation model, i.e., $\boldsymbol{H}_R$ is injective, $\boldsymbol{\varepsilon}_R \in \mathbb{R}^m$, $\boldsymbol{\varepsilon}_R \sim \mathcal{N}(\boldsymbol{\mu}_R, \boldsymbol{R})$ an unknown additive Gaussian noise, with known covariance matrix $\boldsymbol{R}$ but unknown mean $\boldsymbol{\mu}_R$.

By left-multiplying equation (1) by $\boldsymbol{R}^{-1/2}$ ($\boldsymbol{R}^{1/2} \in \mathcal{M}_m(\mathbb{R})$ a lower triangular matrix such that $\boldsymbol{R}^{1/2}\boldsymbol{R}^{T/2} = \boldsymbol{R}$), we obtain:

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

with $\boldsymbol{y} = \boldsymbol{R}^{-1/2}\boldsymbol{y}_R, \boldsymbol{H} = \boldsymbol{R}^{-1/2}\boldsymbol{H}_R, \boldsymbol{\varepsilon} = \boldsymbol{R}^{-1/2}\boldsymbol{\varepsilon}_R \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{I}_m)$, where $\boldsymbol{\mu} = \boldsymbol{R}^{-1/2}\boldsymbol{\mu}_R$.

The noise impacting the measurement vector in equation (2) can be written as:

$$\boldsymbol{\varepsilon} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_0 \tag{3}$$

with $\boldsymbol{\varepsilon}_0 \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_m)$.

We are looking for the estimate that best explains the measurement vector $\boldsymbol{y}$ under the hypothesis that $\boldsymbol{\mu} = \boldsymbol{0}$.

We have the well known least-squares estimate:

$$\widehat{\boldsymbol{x}}_{\text{ls}} = \arg\min_{\widehat{\boldsymbol{x}}} ||\boldsymbol{y} - \boldsymbol{H}\widehat{\boldsymbol{x}}||^2 = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{y}. \tag{4}$$

For commodity, we define:

$$\boldsymbol{S} \triangleq (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T \tag{5}$$

so we can write equation (4) as:

$$\widehat{\boldsymbol{x}}_{\text{ls}} = \boldsymbol{S}\boldsymbol{y}. \tag{6}$$

2. Linear system residuals

For any estimator $\widehat{\boldsymbol{x}}$ of $\boldsymbol{x}$, the residual vector is defined as:

$$\widehat{\boldsymbol{z}} \triangleq \boldsymbol{y} - \boldsymbol{H}\widehat{\boldsymbol{x}} \tag{7}$$

with $\widehat{\boldsymbol{x}}$ an estimate of $\boldsymbol{x}$. By definition, the least-squares estimate $\widehat{\boldsymbol{x}}_{\text{ls}}$ is the one that minimizes the euclidean norm of $\widehat{\boldsymbol{z}}$ in equation (7). The corresponding residual vector is:

$$\widehat{\boldsymbol{z}}_{\text{ls}} = \boldsymbol{y} - \boldsymbol{H}\widehat{\boldsymbol{x}}_{\text{ls}}. \tag{8}$$

Injecting equation (6) into equation (8) and using equation (2), one obtains:

$$\widehat{\boldsymbol{z}}_{\text{ls}} = (\boldsymbol{I}_m - \boldsymbol{H}\boldsymbol{S})\boldsymbol{y} = (\boldsymbol{I}_m - \boldsymbol{H}\boldsymbol{S})\boldsymbol{\varepsilon} \tag{9}$$

and by left-multiplying equation (6) by $\boldsymbol{H}$:

$$\boldsymbol{H}\widehat{\boldsymbol{x}}_{\text{ls}} = \boldsymbol{H}\boldsymbol{S}\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{H}\boldsymbol{S}\boldsymbol{\varepsilon}. \tag{10}$$

For commodity, we define:

$$\boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H})} \triangleq \boldsymbol{H}\boldsymbol{S}, \;\; \boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H})^\perp} \triangleq \boldsymbol{I}_m - \boldsymbol{H}\boldsymbol{S}. \tag{11}$$

One can show that $\boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H})}$ is an orthogonal projector onto $\text{Im}(\boldsymbol{H})$, and that $\boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H})^\perp}$ is an orthogonal projector onto $\text{Im}(\boldsymbol{H})^\perp$, see Figure 1 for a visualization.

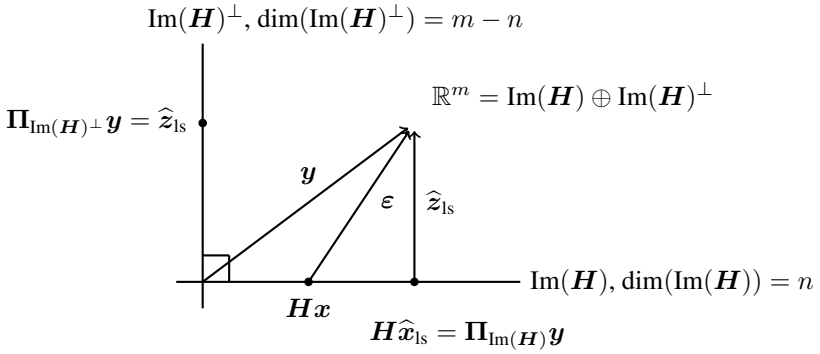

Figure 1: The measurement space $\mathbb{R}^m$ is the direct sum of two orthogonal vector spaces: $\text{Im}(\boldsymbol{H})$ and $\text{Im}(\boldsymbol{H})^\perp$, the least-squares estimate is obtained by an orthogonal projection of $\boldsymbol{y}$ onto $\text{Im}(\boldsymbol{H})$, the residual vector $\widehat{\boldsymbol{z}}_{\text{ls}}$ is obtained by an orthogonal projection of $\boldsymbol{y}$ onto $\text{Im}(\boldsymbol{H})^\perp$.

We define the squared norm of the residual vector:

$$z \triangleq \widehat{\boldsymbol{z}}_{\text{ls}}^T\widehat{\boldsymbol{z}}_{\text{ls}} = \boldsymbol{\varepsilon}^T\boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H})^\perp}^T\boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H})^\perp}\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^T\boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H})^\perp}\boldsymbol{\varepsilon}. \tag{12}$$

If $\boldsymbol{\mu} = \boldsymbol{0}$, Cochran's theorem [Potter and Suman, 1977] states that z follows a central chi-squared distribution (from equation (9)):

$$z \sim \chi^2_{m-n}(0) \tag{13}$$

with $\chi^2_{m-n}(0)$ the chi-squared distribution with $m - n$ degrees of freedom.

If $\boldsymbol{\mu} \neq \boldsymbol{0}$, then z follows a noncentral chi-squared distribution (from equations (9) and (3)):

$$z \sim \chi^2_{m-n}(\lambda), \text{ with } \lambda = \boldsymbol{\mu}^T\boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H})^\perp}\boldsymbol{\mu} \tag{14}$$

with $\chi^2_{m-n}(\lambda)$ the noncentral chi-squared distribution with $m - n$ degrees of freedom and noncentrality parameter λ.

III. FAULT DETECTION AND IDENTIFICATION

When computing the least-squares estimate, it is assumed that $\boldsymbol{\mu} = \mathbf{0}$. The aim of the following sections is to test this hypothesis. Often, we will refer to this quantity as a vector of biases, or bias vector, as these biases are unknown deterministic quantities that impact the measurement vector, and that are not related to the true state that we try to estimate, so it will affect the quality of the estimation of $\boldsymbol{x}$. We will try to identify such a bias vector, and exclude it to obtain the best estimate of $\boldsymbol{x}$ that is achievable. A two step process is followed: step 1 consists of deciding whether or not there is a bias in the measurement vector, and if so, step 2 two will identify the biased measurement to be able to exclude it from the solution.

1. Hypothesis/boundaries

In the case $\boldsymbol{\mu} = \mathbf{0}$, then equation (13) holds. If $\boldsymbol{\mu} \neq \mathbf{0}$, then we omitted the mean of $\boldsymbol{\varepsilon}$, equation (14) holds. These two equations imply that to detect a bias, we have the necessary condition: $m - n > 0$.

We do not know the bias that actually impacts the measurement vector, so we cannot compute λ in equation (14). In our application, we will define a worst-case bias $\boldsymbol{\mu}_{\text{wc}}$:

- we will assume that a single component of $\boldsymbol{\mu}_{\text{wc}}$ is non-null,
- the amplitude of this bias lives in a discrete set, so we define the smallest amplitude, in absolute value, of this set as μ_{wc},
- it happens on worst measurement, i.e., on component of $\boldsymbol{\mu}$ such that the noncentrality parameter λ in equation (14) is minimum.

Therefore, we can write:

$$\boldsymbol{\mu}_{\text{wc}} = \mu_{\text{wc}} \boldsymbol{e}_{\text{wc}} \tag{15}$$

with $\boldsymbol{e}_{\text{wc}}$ a vector of the canonical basis of $\mathbb{R}^m$, and the worst-case noncentrality parameter is defined:

$$\lambda_{\text{wc}} \triangleq \boldsymbol{\mu}_{\text{wc}}^T \boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H})^\perp} \boldsymbol{\mu}_{\text{wc}}. \tag{16}$$

2. Single fault detection

The first test is a binary hypothesis testing, deciding whether the bias is null or not. To answer it, we write the two hypotheses (the upper index indicates step 1 of the detection and exclusion process) as:

$$\begin{cases} H_0^{(1)} : & \boldsymbol{\mu} = \mathbf{0} \\ H_1^{(1)} : & \boldsymbol{\mu} \neq \mathbf{0}. \end{cases} \tag{17}$$

We will predict that $H_1^{(1)}$ holds if the residual becomes too large, that is if [Kay, 2009]:

$$z \geq T \tag{18}$$

with T a decision threshold and z as in equation (12). We will set T by a false-alarm constraint, making T depending only on the distribution equation (13) and on a manually set probability of false alarm.

Most of the time, the true bias $\boldsymbol{\mu}$ will be different from $\boldsymbol{\mu}_{\text{wc}}$, but if it contains at most a single fault, under the conditions on $\boldsymbol{\mu}_{\text{wc}}$, the true noncentrality parameter λ will be larger than λ_{wc}, so the test will deliver a lower bound on probability of detection.

If we made the decision $H_1^{(1)}$, the second step is to identify which component of $\boldsymbol{\mu}$ is non-null.

3. Fault identification

If we made a decision that a bias is present (i.e., we decided $H_1^{(1)}$ in the previous step), we can try to identify the fault.

We denote $\boldsymbol{e}_i$ the i^{th} vector of the canonical basis of $\mathbb{R}^m$, $i \in \{1, \ldots, m\}$. We made the assumption that a single fault is present at most, let us assume that it affects component j with an amplitude μ:

$$\boldsymbol{\mu} = \mu \boldsymbol{e}_j. \tag{19}$$

Then, injecting equations (3) and (19) into equation (2), we can rewrite the measurement model as:

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{H} & \boldsymbol{e}_j \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \mu \end{pmatrix} + \boldsymbol{\varepsilon}_0. \tag{20}$$

For commodity, we define for any i:

$$\boldsymbol{H}_i \triangleq (\boldsymbol{H} \quad \boldsymbol{e}_i) \tag{21}$$

so equation (20) becomes:

$$\boldsymbol{y} = \boldsymbol{H}_j \begin{pmatrix} \boldsymbol{x} \\ \mu \end{pmatrix} + \boldsymbol{\varepsilon}_0. \tag{22}$$

It is important to remark that we can treat equation (22) as a measurement model if and only if $\boldsymbol{H}_j$ is injective, i.e., if it is of rank $n+1$. The same applies to all matrices $\boldsymbol{H}_i$. Put another way, it means that if $\boldsymbol{H}$ is full rank and $\boldsymbol{H}_i$ is not, then the i^{th} measurement cannot be discarded to estimate $\boldsymbol{x}$.

As we do not know j, we will build m new hypotheses $H_i^{(2)},\ i \in \{1, \ldots, m\}$:

$$H_i^{(2)} : i = j \tag{23}$$

and test each one of them. The hypothesis $H_i^{(2)}$ can be interpreted as measurement i being faulty. The upper index indicates step 2 of the detection and identification process.

For commodity, we define:

$$\boldsymbol{S}_i \triangleq (\boldsymbol{H}_i^T \boldsymbol{H}_i)^{-1} \boldsymbol{H}_i^T. \tag{24}$$

Under hypothesis $H_i^{(2)}$, we can compute a least-squares estimate:

$$\begin{pmatrix} \widehat{\boldsymbol{x}}_{\text{ls}_i} \\ \widehat{\mu}_i \end{pmatrix} = \boldsymbol{S}_i \boldsymbol{y} \tag{25}$$

with $\boldsymbol{y}$ as in equation (22), and build a residual vector as:

$$\widehat{\boldsymbol{z}}_{\text{ls}_i} = \boldsymbol{y} - \boldsymbol{H}_i \begin{pmatrix} \widehat{\boldsymbol{x}}_{\text{ls}_i} \\ \hat{\mu}_i \end{pmatrix} \tag{26}$$

and we define:

$$z_i \triangleq \widehat{\boldsymbol{z}}_{\text{ls}_i}^T \widehat{\boldsymbol{z}}_{\text{ls}_i}. \tag{27}$$

In Appendix A we show that:

- if $i = j$, both $\widehat{\boldsymbol{x}}_{\text{ls}_i}$ and $\widehat{\mu}$ will be centered on the hidden states, hence z_i follows a central chi-squared distribution with $m - n - 1$ degrees of freedom,
- if $i \neq j$, the estimates will not be centered on the hidden states, a bias will remain in the residual vector, hence z_i follows a noncentral chi-squared distribution with $m - n - 1$ degrees of freedom.

This implies that to identify the bias, we have the necessary condition: $m - n - 1 > 0$.

In the case of $i \neq j$, we will not be able to compute the true noncentrality parameter, since we do not know j nor the hidden parameters and bias. But what we know is that if the model i is not the true model j, then the squared norm of its residual vector does not follow a central chi-squared distribution. From this assessment, we can build a maximum likelihood estimation problem. We will decide i such that:

$$\hat{i} = \underset{i \in \{1,\ldots,m\}}{\arg\max}\ \chi^2_{m-n-1}(0)(z_i). \tag{28}$$

The model with the correct vector $\boldsymbol{e}_i$, i.e., with $i = j$, will generate a residual vector whose squared norm follows a central chi-squared distribution, i.e., it will statistically have the highest likelihood of being central. For the other models, a bias will remain in the residual vector, and the likelihood of the residual squared norm being central is statistically lower or equal.

It is important to keep in mind that under the proposed hypothesis, in the presence of two or more biases, we can have a noncentrality parameter λ lower than λ_{wc}. Unfortunately, under such conditions, this test does not deliver a lower bound on probability of detection.

Now that we have seen how to detect and identify a single fault in a linear observation model, we can move onto factor graph, as they will handle the case of nonlinear observation models.

IV. FACTOR GRAPH OPTIMIZATION

A factor graph is a bipartite graph, with two set of nodes, the first one being states to estimate, called variables, the second one being what we call factors. A factor contains a measurement vector, an observation model, and a noise model. Each factor is connected to the subset of variables appearing in its observation model. The aim of a factor graph is to represent an estimation problem.

For K factors in the factor graph, $k \in \{1, \dots, K\}$, we have the observation model of factor k:

$$\boldsymbol{y}_k = h_k(\boldsymbol{x}) + \boldsymbol{\varepsilon}_k \tag{29}$$

with $\boldsymbol{x} \in \mathbb{R}^n$ a hidden state vector, $\boldsymbol{y}_k \in \mathbb{R}^{m_k}$ a known measurement vector, $h_k : \mathbb{R}^n \to \mathbb{R}^{m_k}$ a known (possibly nonlinear) model, $\boldsymbol{\varepsilon}_k \in \mathbb{R}^{m_k}, \boldsymbol{\varepsilon}_k \sim \mathcal{N}(\mathbf{0}_{m_k,1}, \boldsymbol{R}_k)$ an unknown additive Gaussian noise, with known covariance matrix $\boldsymbol{R}_k$. From equation (29), the conditional probability density function of observing $\boldsymbol{y}_k$ given $\boldsymbol{x}$ is:

$$p_k(\boldsymbol{y}_k|\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^{m_k}|\boldsymbol{R}_k|}} \exp(-\frac{1}{2}||\boldsymbol{y}_k - h_k(\boldsymbol{x})||^2_{\boldsymbol{R}_k}) \tag{30}$$

with the Mahalanobis norm:

$$||\boldsymbol{v}||^2_{\boldsymbol{\Omega}} \triangleq \boldsymbol{v}^T \boldsymbol{\Omega}^{-1} \boldsymbol{v}. \tag{31}$$

The likelihood is defined as any function proportional to $p_k(\boldsymbol{y}_k|\widehat{\boldsymbol{x}})$, so we define the likelihood:

$$l_k(\widehat{\boldsymbol{x}}; \boldsymbol{y}_k) \triangleq \exp(-\frac{1}{2}||\boldsymbol{y}_k - h_k(\widehat{\boldsymbol{x}})||^2_{\boldsymbol{R}_k}) \propto p_k(\boldsymbol{y}_k|\widehat{\boldsymbol{x}}) \tag{32}$$

with the notation indicating that the likelihood is not a function of $\boldsymbol{y}_k$, which acts as a parameter here, but a function of $\widehat{\boldsymbol{x}}$. We define the residual vector:

$$\boldsymbol{z}_k(\widehat{\boldsymbol{x}}) \triangleq \boldsymbol{y}_k - h_k(\widehat{\boldsymbol{x}}) \tag{33}$$

so the likelihood (32) becomes:

$$l_k(\widehat{\boldsymbol{x}}; \boldsymbol{y}_k) = \exp(-\frac{1}{2}||\boldsymbol{z}_k(\widehat{\boldsymbol{x}})||^2_{\boldsymbol{R}_k}). \tag{34}$$

We can vertically stack all the measurements:

$$\boldsymbol{y} \triangleq \begin{pmatrix} \boldsymbol{y}_1^T & \dots & \boldsymbol{y}_K^T \end{pmatrix}^T \in \mathbb{R}^m, \; m = \sum_{k=1}^{K} m_k. \tag{35}$$

If the noises impacting the K factors are all independent, i.e., if $\forall (k, k'), k \neq k'$ we have $\mathbb{E}[\boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_{k'}^T] = \mathbf{0}_{m_k, m_{k'}}$, $p_k(\boldsymbol{y}_k|\boldsymbol{x})$ and $p_{k'}(\boldsymbol{y}_{k'}|\boldsymbol{x})$ are independent, so:

$$p(\boldsymbol{y}|\boldsymbol{x}) = \prod_{k=1}^{K} p_k(\boldsymbol{y}_k|\boldsymbol{x}) \tag{36}$$

and the likelihood $l(\widehat{\boldsymbol{x}}; \boldsymbol{y})$ is the product of the likelihoods:

$$l(\widehat{\boldsymbol{x}}; \boldsymbol{y}) = \prod_{k=1}^{K} l_k(\widehat{\boldsymbol{x}}; \boldsymbol{y}_k). \tag{37}$$

From this, we can define the maximum likelihood vector $\widehat{\boldsymbol{x}}^*$:

$$\widehat{\boldsymbol{x}}^* \triangleq \arg\max_{\widehat{\boldsymbol{x}}} l(\widehat{\boldsymbol{x}}; \boldsymbol{y}). \tag{38}$$

Solving the optimization problem equation (38) is equivalent to minimizing its negative log, i.e.,

$$\arg\max_{\widehat{\boldsymbol{x}}} l(\widehat{\boldsymbol{x}}; \boldsymbol{y}) = \arg\min_{\widehat{\boldsymbol{x}}} - \log l(\widehat{\boldsymbol{x}}; \boldsymbol{y}) = \arg\min_{\widehat{\boldsymbol{x}}} \sum_{k=1}^{K} ||\boldsymbol{z}_k(\widehat{\boldsymbol{x}})||^2_{\boldsymbol{R}_k} \tag{39}$$

which is a nonlinear least-squares problem. To solve such a problem, we can use the Gauss-Newton algorithm: it is an iterative solver, i.e., it starts from an initial point $\boldsymbol{x}_0$, finds a step $\Delta\boldsymbol{x}^*$, applies it, and iterates until a convergence criterion is met:

$$\Delta\boldsymbol{x}^* = \underset{\Delta\boldsymbol{x}}{\arg\min} \sum_{k=1}^{K} ||\boldsymbol{z}_k(\boldsymbol{x}_0 + \Delta\boldsymbol{x})||^2_{\boldsymbol{R}_k} \tag{40}$$

$$= \underset{\Delta\boldsymbol{x}}{\arg\min} \sum_{k=1}^{K} ||\boldsymbol{R}_k^{-1/2}\boldsymbol{z}_k(\boldsymbol{x}_0 + \Delta\boldsymbol{x})||^2. \tag{41}$$

To do so, the Gauss-Newton approximation consists of writing the first-order Taylor development of the observation model h_k in equation (33). Equation (40) becomes:

$$\Delta\boldsymbol{x}^* = \underset{\Delta\boldsymbol{x}}{\arg\min} \sum_{k=1}^{K} ||\boldsymbol{z}_k(\boldsymbol{x}_0) - \boldsymbol{J}_k|_{\boldsymbol{x}_0}\Delta\boldsymbol{x}||^2_{\boldsymbol{R}_k} \tag{42}$$

$$= \underset{\Delta\boldsymbol{x}}{\arg\min} \sum_{k=1}^{K} ||\boldsymbol{R}_k^{-1/2}\boldsymbol{z}_k(\boldsymbol{x}_0) - \boldsymbol{R}_k^{-1/2}\boldsymbol{J}_k|_{\boldsymbol{x}_0}\Delta\boldsymbol{x}||^2 \tag{43}$$

with $\boldsymbol{J}_k|_{\boldsymbol{x}_0}$ the Jacobian matrix of h_k computed at $\boldsymbol{x}_0$. To get rid of the sum in equation (43), we can vertically stack all the whitened residuals:

$$\boldsymbol{z}(\boldsymbol{x}_0) \triangleq \begin{pmatrix} (\boldsymbol{R}_1^{-1/2}\boldsymbol{z}_1(\boldsymbol{x}_0))^T & \dots & (\boldsymbol{R}_K^{-1/2}\boldsymbol{z}_K(\boldsymbol{x}_0))^T \end{pmatrix}^T \tag{44}$$

and do the same to the Jacobian matrices:

$$\boldsymbol{J}|_{\boldsymbol{x_0}} \triangleq \begin{pmatrix} (\boldsymbol{R}_1^{-1/2}\boldsymbol{J}_1|_{\boldsymbol{x}_0})^T & \dots & (\boldsymbol{R}_K^{-1/2}\boldsymbol{J}_K|_{\boldsymbol{x}_0})^T \end{pmatrix}^T \tag{45}$$

so that equation (43) becomes:

$$\Delta\boldsymbol{x}^* = \underset{\Delta\boldsymbol{x}}{\arg\min} ||\boldsymbol{z}(\boldsymbol{x}_0) - \boldsymbol{J}|_{\boldsymbol{x_0}}\Delta\boldsymbol{x}||^2. \tag{46}$$

The solution to this optimization problem is given in equation (4). Once this problem is solved, we apply the update:

$$\boldsymbol{x}_0 \leftarrow \boldsymbol{x}_0 + \alpha\Delta\boldsymbol{x}^* \tag{47}$$

with α a step size hyperparameter. In the linear case presented in the previous section, we had a condition on the rank of our observation model. Here, the observations models of the factors might be nonlinear, so in order to converge, we will iterate until the convergence criterion is met, which will always be the case if the h_k models are Lipschitz continuously differentiable, if the Jacobian matrices are uniformly full rank, and if the Armijo and Wolfe step size conditions are met [Nocedal and Wright, 2006].

Equation (46) has the same form as equation (4), so once at convergence ($\Delta\boldsymbol{x}^* \approx \boldsymbol{0}$), we will apply the detection theory that we saw in the previous section to the current problem.

V. APPLICATION TO GNSS

We are going to build a factor graph containing carrier phase observations, as we target precise positioning applications.

1. Carrier phase observation model

In order to reach centimeter level positioning, we can use GNSS carrier phase measurements, as these measurements are impacted by a very small random noise. The drawback of these measurements is that they are ambiguous, i.e., there is an unknown number of wavelengths between our receiver and a given satellite. This number is called the ambiguity. In theory, if a receiver is tracking a satellite, then this ambiguity should stay constant. However, the phase tracking process can converge to a different stable tracking point, which will result in a change in the ambiguity. This type of event is called a cycle slip, and can be due to perturbation of the signal reception due to excessive noise (low elevation satellite, interference), excessive signal dynamics (dynamic maneuver of the receiver) or propagation impairments (multipath). A receiver at epoch t of coordinates $\boldsymbol{x}_t = \begin{pmatrix} x_t & y_t & z_t \end{pmatrix}^T$ is tracking a satellite with an identifier q and of coordinates $\boldsymbol{x}_t^q = \begin{pmatrix} x_t^q & y_t^q & z_t^q \end{pmatrix}^T$, we denote $\boldsymbol{r}_t^q = \boldsymbol{x}_t^q - \boldsymbol{x}_t$ the vector receiver to satellite q, at epoch t. The carrier phase measurement used in this experiment is the GPS L1 (a wavelength of 19 cm), and is modeled as follows:

$$\lambda_{\text{L1}}\tilde{\Phi}_t^q = ||\boldsymbol{r}_t^q|| + c(dt_t - dt_t^q) + \lambda_{\text{L1}}N_t^q - I_t^q + T_t^q + \delta^q_{\text{eph}_t} + m_t^q + \tilde{\varepsilon}_t^q \tag{48}$$

where $\tilde{\Phi}_t^q$ is the carrier phase raw measurement of satellite q at epoch t, λ_{L1} is the wavelength of the GPS L1 band, c is the speed of light, dt_t the receiver clock bias with respect to GPS time at epoch t, dt_t^q the satellite clock bias with respect to GPS time at epoch t, I_t^q the ionospheric delay at epoch t, T_t^q the tropospheric delay at epoch t, $\delta^q_{\text{eph}_t}$ the ephemeris error of satellite q at epoch t, m_t^q the multipath error at epoch t, $N_t^q \in \mathbb{Z}$ the integer ambiguity, and ε_t^q a centered Gaussian noise.

After correction of the ionospheric delay, tropospheric delay, satellite and receiver clock biases, ephemeris and multipath errors [Sanz et al., 2013], the model becomes:

$$\lambda_{\text{L1}}\Phi_t^q = ||\boldsymbol{r}_t^q|| + \lambda_{\text{L1}}N_t^q + \varepsilon_t^q \tag{49}$$

with $\varepsilon_t^q \sim \mathcal{N}(0, {\sigma^q_{\Phi_t}}^2)$ the error made of all the residuals errors after correction.

If we assume that, for all satellite tracked during a time window of w epochs, the ambiguity remains constant at all epochs ($N_1^q = N_2^q = \ldots = N_w^q = N^q$), we can model the observations by the graph in Figure 2.

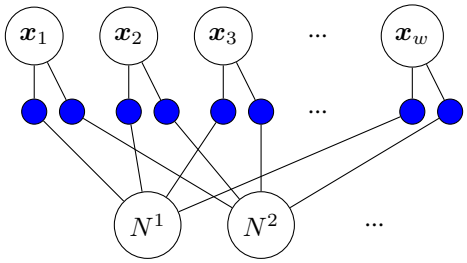

Figure 2: A factor graph made of corrected carrier phase measurements. Each ambiguity N corresponds to a satellite, each $\boldsymbol{x}_t$ corresponds to a 3D position at a given epoch t.

If a cycle slip happens, then all the carrier phase observations after that cycle slip will be faulty. This leads to more than one fault. Our theory does not provide results on multiple faults, so we need to transform our graph in a way to be in the single fault case.

2. Consecutive time difference carrier phase

For a satellite q tracked through all epochs from 1 to w, we denote $\boldsymbol{\Phi}^q$ the vector of all carrier phase observations of this satellite:

$$\lambda_{\text{L1}}\boldsymbol{\Phi}^q = \begin{pmatrix} \lambda_{\text{L1}}\Phi_1^q \\ \vdots \\ \lambda_{\text{L1}}\Phi_w^q \end{pmatrix} = \begin{pmatrix} ||\boldsymbol{r}_1^q|| + \lambda_{\text{L1}}N_1^q \\ \vdots \\ ||\boldsymbol{r}_w^q|| + \lambda_{\text{L1}}N_w^q \end{pmatrix} + \underbrace{\begin{pmatrix} \varepsilon_1^q \\ \vdots \\ \varepsilon_w^q \end{pmatrix}}_{\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{R}^q)} \tag{50}$$

with $\boldsymbol{R}^q = \begin{pmatrix} {\sigma^q_{\Phi_1}}^2 & & \\ & \ddots & \\ & & {\sigma^q_{\Phi_w}}^2 \end{pmatrix}$ a diagonal matrix. This is the reason why Figure 2 is made of a lot of sparse scalar factors.

In the case of a single cycle slip, one way to be in the single fault case is to compute time differences between consecutive observations. To do so, we multiply equation (50) by a differentiation matrix:

$$\boldsymbol{D} = \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix} \tag{51}$$

leading to:

$$\text{TDCP}^q = \boldsymbol{D}\lambda_{\text{L1}}\boldsymbol{\Phi}^q = \begin{pmatrix} ||\boldsymbol{r}_2^q|| - ||\boldsymbol{r}_1^q|| + \lambda_{\text{L1}}(N_2^q - N_1^q) \\ \vdots \\ ||\boldsymbol{r}_w^q|| - ||\boldsymbol{r}_{w-1}^q|| + \lambda_{\text{L1}}(N_w^q - N_{w-1}^q) \end{pmatrix} + \underbrace{\boldsymbol{D}\begin{pmatrix} \varepsilon_1^q \\ \vdots \\ \varepsilon_w^q \end{pmatrix}}_{\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{D}\boldsymbol{R}^q\boldsymbol{D}^T)} . \tag{52}$$

If we assume constant ambiguity, then the system becomes:

$$\text{TDCP}^q = \begin{pmatrix} ||\boldsymbol{r}_2^q|| - ||\boldsymbol{r}_1^q|| \\ \vdots \\ ||\boldsymbol{r}_w^q|| - ||\boldsymbol{r}_{w-1}^q|| \end{pmatrix} + \underbrace{\boldsymbol{D} \begin{pmatrix} \varepsilon_1^q \\ \vdots \\ \varepsilon_w^q \end{pmatrix}}_{\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{D}\boldsymbol{R}^q\boldsymbol{D}^T)} . \tag{53}$$

We can represent the following system with the factor graph in Figure 3. The reader can check that the matrix $\boldsymbol{D}\boldsymbol{R}^q\boldsymbol{D}^T$ is not diagonal, there are correlations between TDCPs. Therefore, the factors are connected to all the variables.

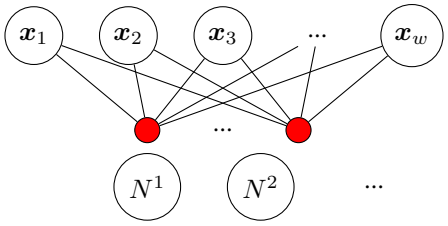

Figure 3: A factor graph made of TDCP measurements. The factors are connected to all the variables, excepted the ambiguities, indicating correlation between scalar TDCPs.

In the case of a single cycle slip, only the TDCP with the two different ambiguities will be faulty. Indeed, as the ambiguities are not connected to any factor, they are not observable, which goes against observability/injectivity conditions required in the theory. But as they do not appear in the observation model, we are in the trivial case where we can simply remove them from the graph, thereby recovering observability for the remaining system.

3. Experiments

To test the proposed application, we built a simulator, where we used a fixed receiver, precise ephemeris of the GPS constellation, with a period of 5 minutes between epochs. In order to have a ground truth regarding cycle slips, we generated carrier phase measurement from the receiver, following equation (49). The noise level of these observations was set to $\sigma^q_{\Phi_t} = 5$ cm, $\forall t, \forall q$, and with no correlation between carrier phase measurements, as in Figure 2. For programming reasons, for a graph of TDCP measurements of length w epochs, we kept in the graph the subset of satellites that are visible from epoch 1 through w, without discontinuity. w ranges from 2 to the highest value possible, until the graph becomes unobservable, due to lack of enough observations. So as w increases, the number of satellites kept will decrease. From these carrier phase observations, we built the TDCP measurements.

To test our cycle slip detector, we generated 300 graphs for each value of w, with a probability of 0.5 to have exactly one cycle slip, of smallest amplitude ($N^q_{t+1} - N^q_t = \pm 1$). This cycle slip can happen on any satellite and at any epoch. Thanks to our hypothesis on the noncentrality parameter, the empirical probability of detection should be higher than the lower bound. The lower bound of probability of detection is given by the theoretical development. The probability of false alarm was set to a value of 0.05. The results are given in Figure 4.

What comes out from these experiments is that if we consider a constant number of visible satellites, for example for $w \in \{2, \dots, 8\}$, excepted when we move from 2 to 3 epochs, the lower bound of probability of detection decreases, so the worst case becomes worst as w increases. On the other hand, the number of factor increases as w increases, so in our experiments, the probability of having a cycle slip on the worst TDCP decreases as w increases, hence a higher empirical probability of detection, that is increasing with w.

Then, we did the same experiment, but this time, with a lower noise level: $\sigma^q_{\Phi_t} = 2$ cm, $\forall t, \forall q$. Compared to the previous experiment, the noncentrality parameter equation (16) is increased, leading to a higher probability of detection, as in Figure 5.

The improved detection indicates that the better the correction of atmospheric errors, clock biases, and all the other terms appearing in equation (48), the better the detection capability.

The motivation behind the use of synthetic measurements is that we have a ground truth regarding cycle slips. Moreover, the use of real measurements would not have changed the lower bound of probability of detection, as it only depends on the geometry, not on the actual measurements, under the proposed hypotheses.

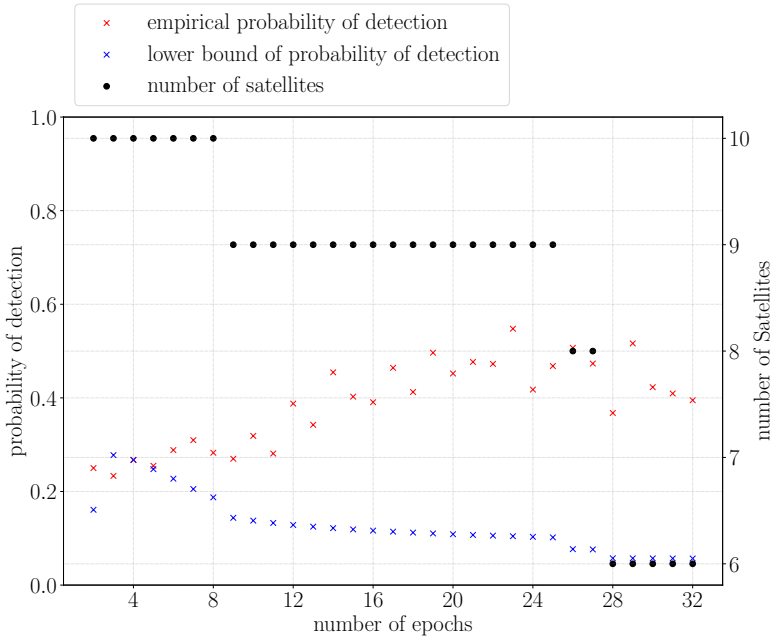

Figure 4: Probability of detection of single cycle slip using a factor graph with TDCP factors, with $\sigma^q_{\Phi_t} = 5$ cm, $\forall t, \forall q$. Bottom axis: number of epochs (w), left axis: probabilities of detection, right axis: number of visible satellites throughout all epochs. The lower bound of probability of detection is the probability of having the squared norm of the residual vector greater to the threshold T when there is a cycle slip of one wavelength on the worst factor.

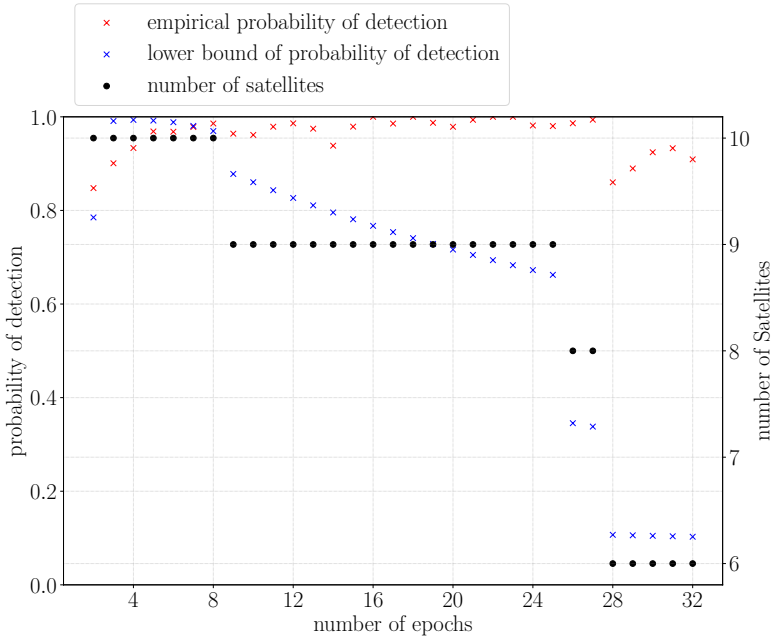

Figure 5: Probability of detection of single cycle slip using a factor graph with TDCP factors, with $\sigma^q_{\Phi_t} = 2$ cm, $\forall t, \forall q$. Bottom axis: number of epochs (w), left axis: probabilities of detection, right axis: number of visible satellites throughout all epochs. The lower bound of probability of detection is the probability of having the squared norm of the residual vector greater to the threshold T when there is a cycle slip of one wavelength on the worst factor.

CONCLUSION

We have proposed a very general theory of least-squares estimation under linear and nonlinear models with an additive Gaussian perturbation. Then, we provided a geometric approach of this theory, and wrote statistics of the squared norm of the residual, in the case of faults on the measurements. We proposed to investigate a set of possible faults, and from this, provided statistics of detection. This theory allows us to deliver a lower bound of probability of detection of single faults living in a discrete set, given a probability of false alarm. Then, we built an optimization problem over a finite set to identify the fault. We then moved onto factor graph optimization, and wrote it in a way that allows us to use the fault detection and identification theory to problems that can be modelled by a factor graph. We then transposed our general theory to cycle slip detection and identification, as we target precise GNSS positioning applications. We proposed a graph structure that fits our theory, so we were able to deliver a lower bound of probability of detection of a single cycle slip of smallest amplitude. We saw that the lowest probability of detection does not necessarily increases as the degree of freedom of the model increases. The reader must keep in mind that in the case of multiple faults, our theory does not provide results, nor if the fault does not live in a discrete set. Multiple studies could be carried, about multiple faults living in a discrete set, about the use of prior information on cycle slip for a better detection and identification of fault.

REFERENCES

[Bai et al., 2022] Bai, X., Wen, W., and Hsu, L.-T. (2022). Time-Correlated Window-Carrier-Phase-Aided GNSS Positioning Using Factor Graph Optimization for Urban Positioning. *IEEE Transactions on Aerospace and Electronic Systems*, 58(4):3370–3384.

[Congram and Barfoot, 2022] Congram, B. and Barfoot, T. (2022). Field Testing and Evaluation of Single-Receiver GPS Odometry for Use in Robotic Navigation. *Field Robotics*, 2(1):1849–1873.

[Farooq et al., 2018] Farooq, S. Z., Yang, D., Jin, T., and Ada, E. N. J. (2018). Survey of Cycle Slip Detection & Correction Techniques for Single Frequency Receivers. In *2018 IEEE 18th International Conference on Communication Technology (ICCT)*, pages 957–961, Chongqing. IEEE.

[Hajiyev, 2016] Hajiyev, C. (2016). An Innovation Approach Based Sensor Fault Detection and Isolation. *IFAC-PapersOnLine*, 49(17):420–425.

[Hofmann-Wellenhof et al., 2008] Hofmann-Wellenhof, B., Lichtenegger, H., and Wasle, E. (2008). *GNSS - Global Navigation Satellite Systems: GPS, GLONASS, Galileo and more*. Springer, Wien New York.

[Joerger et al., 2014] Joerger, M., Chan, F.-C., and Pervan, B. (2014). Solution Separation Versus Residual-Based RAIM: Solution Separation Versus Residual-Based RAIM. *Navigation*, 61(4):273–291.

[Kay, 2009] Kay, S. M. (2009). *Fundamentals of statistical processing, Volume 2: Detection theory*. Pearson Education India.

[Kerhascoet et al., 2017] Kerhascoet, H., Merien, P., Laurent, J., Senn, E., and Hauville, F. (2017). Sensor Fault Detection and Signal Improvement using Predictive Filters. *International Journal of Small Craft Technology*, 159(Part B1):186.

[Kirkko-Jaakkola et al., 2009] Kirkko-Jaakkola, M., Traugott, J., Odijk, D., Collin, J., Sachs, G., and Holzapfel, F. (2009). A raim approach to GNSS outlier and cycle slip detection using L1 carrier phase time-differences. In *2009 IEEE Workshop on Signal Processing Systems*, pages 273–278, Tampere, Finland. IEEE.

[Knowles and Gao, 2023] Knowles, D. and Gao, G. (2023). Euclidean Distance Matrix-Based Rapid Fault Detection and Exclusion. *NAVIGATION: Journal of the Institute of Navigation*, 70(1):navi.555.

[Langley, 2014] Langley, R. B. (2014). Innovation: Cycle Slips - GPS World.

[Li and Melachroinos, 2019] Li, T. and Melachroinos, S. (2019). An enhanced cycle slip repair algorithm for real-time multi-GNSS, multi-frequency data processing. *GPS Solutions*, 23(1):1.

[Lotfy et al., 2022] Lotfy, A., Abdelfatah, M., and El-Fiky, G. (2022). Improving the performance of GNSS precise point positioning by developed robust adaptive Kalman filter. *The Egyptian Journal of Remote Sensing and Space Science*, 25:919–928.

[Nocedal and Wright, 2006] Nocedal, J. and Wright, S. J. (2006). *Numerical optimization*. Springer series in operations research. Springer, New York, 2nd ed edition.

[Potter and Suman, 1977] Potter, J. and Suman, M. (1977). Thresholdless redundancy management with arrays of skewed instruments. *1977. 25*.

[Roberts, 2019] Roberts, G. W. (2019). Noise comparison of triple frequency GNSS carrier phase, doppler and pseudorange observables. *Measurement*, 144:328–344.

[Sanz et al., 2013] Sanz, J., Juan, J., and Hernández-Pajares, M. (2013). Gnss data processing, vol. i: Fundamentals and algorithms. noordwijk, the netherlands: Esa communications. Technical report, ESTEC TM-23/1, 2013.

[Wang et al., 2022] Wang, Z., Hou, X., Dan, Z., and Fang, K. (2022). Adaptive Kalman filter based on integer ambiguity validation in moving base RTK. *GPS Solutions*, 27(1):34.

[Wen and Hsu, 2021] Wen, W. and Hsu, L.-T. (2021). Towards Robust GNSS Positioning and Real-time Kinematic Using Factor Graph Optimization. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5884–5890.

[Wen and Hsu, 2022] Wen, W. and Hsu, L.-T. (2022). Factor Graph Optimization for Tightly-Coupled GNSS Pseudorange/Doppler/Carrier Phase/INS Integration: Performance in Urban Canyons of Hong Kong. pages 2178–2189.

A. MULTIPLE MODELS AND MULTIPLE HYPOTHESIS

We remind that the true measurement model follows hypothesis $H_j^{(2)}$.

If $i = j$, then equation (25) becomes:

$$\begin{pmatrix} \widehat{\boldsymbol{x}}_{\text{ls}_i} \\ \hat{\mu}_i \end{pmatrix} = \boldsymbol{S}_i(\boldsymbol{H}_i \begin{pmatrix} \boldsymbol{x} \\ \mu \end{pmatrix} + \boldsymbol{\varepsilon}_0) = \begin{pmatrix} \boldsymbol{x} \\ \mu \end{pmatrix} + \boldsymbol{S}_i\boldsymbol{\varepsilon}_0 \tag{54}$$

so the estimate is centered on the true parameters, since $\boldsymbol{\varepsilon}_0$ is of null mean. Still in the case $i = j$, we can write the residual vector equation (26) of this estimate:

$$\widehat{\boldsymbol{z}}_{\text{ls}_i} = \boldsymbol{y} - \boldsymbol{H}_i \begin{pmatrix} \widehat{\boldsymbol{x}}_{\text{ls}_i} \\ \hat{\mu}_i \end{pmatrix} = (\boldsymbol{I}_m - \boldsymbol{H}_i\boldsymbol{S}_i)\boldsymbol{\varepsilon}_0 = \boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H}_i)^\perp}\boldsymbol{\varepsilon}_0. \tag{55}$$

The state vector that we estimate is of dimension $n + 1$, so $\text{Im}(\boldsymbol{H}_i)^\perp$ is of dimension $m - n - 1$, the residual vector is of null mean, so the residual squared norm will follow a central chi-squared distribution with $m - n - 1$ degrees of freedom.

On the other hand, if $i \neq j$, then equation (25) becomes:

$$\begin{pmatrix} \widehat{\boldsymbol{x}}_{\text{ls}_i} \\ \hat{\mu}_i \end{pmatrix} = \boldsymbol{S}_i\boldsymbol{y} = \boldsymbol{S}_i(\boldsymbol{H}_j \begin{pmatrix} \boldsymbol{x} \\ \mu \end{pmatrix} + \boldsymbol{\varepsilon}_0) \tag{56}$$

and the residual vector equation (26) becomes:

$$\begin{aligned} \widehat{\boldsymbol{z}}_{\text{ls}_i} &= \boldsymbol{y} - \boldsymbol{H}_i \begin{pmatrix} \widehat{\boldsymbol{x}}_{\text{ls}_i} \\ \hat{\mu}_i \end{pmatrix} && (57)\\ &= \boldsymbol{y} - \boldsymbol{H}_i\boldsymbol{S}_i\boldsymbol{y} && (58)\\ &= (\boldsymbol{I}_m - \boldsymbol{H}_i\boldsymbol{S}_i)\boldsymbol{y} && (59)\\ &= (\boldsymbol{I}_m - \boldsymbol{H}_i\boldsymbol{S}_i)(\boldsymbol{H}_j \begin{pmatrix} \boldsymbol{x} \\ \mu \end{pmatrix} + \boldsymbol{\varepsilon}_0). && (60) \end{aligned}$$

Here, the estimate is of mean $\boldsymbol{S}_i\boldsymbol{H}_j \begin{pmatrix} \boldsymbol{x} \\ \mu \end{pmatrix}$, and the residual squared norm follows a noncentral chi-squared distribution, with $m - n - 1$ degrees of freedom and a noncentrality parameter:

$$\lambda_i = \begin{pmatrix} \boldsymbol{x} \\ \mu \end{pmatrix}^T \boldsymbol{H}_j^T\boldsymbol{\Pi}_{\text{Im}(\boldsymbol{H}_i)^\perp}\boldsymbol{H}_j \begin{pmatrix} \boldsymbol{x} \\ \mu \end{pmatrix}. \tag{61}$$