Peer Reviewed

Title:

A New Approach for Computationally Efficient and Reliable Carrier Integer Ambiguity Resolution in GPS/INS

Author:

Chen, Yiming

Acceptance Date:

2014

Series:

UC Riverside Electronic Theses and Dissertations

Degree:

Ph.D., Electrical EngineeringUC Riverside

Advisor(s):

Farrell, Jay

Committee:

Mourikis, Anastasios, Kassas, Zak

Permalink:

http://escholarship.org/uc/item/0579t1x5

Abstract:

UNIVERSITY OF CALIFORNIA
RIVERSIDE

A New Approach for Computationally Efficient and Reliable Carrier Integer
Ambiguity Resolution in GPS/INS

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Yiming Chen

December 2014

Dissertation Committee:

    Professor Jay Farrell, Chairperson
    Professor Anastasios Mourikis
    Professor Zak Kassas

The Dissertation of Yiming Chen is approved:

_____

_____

_____
                                    Committee Chairperson

University of California, Riverside

Acknowledgments

First of all, I would like to give my special thanks to my advisor, Prof. Jay A. Farrell from whom I have learned how to study in graduate school and in my future life, how to do research and be a good engineer. He gave me my birth in academia. Not only his professional knowledge, but also his rigorous thinking and writing style will always inspire me. I really enjoy working with him and expect further collaboration in future.

I want to thank other members of my oral exam and defense committee, Prof. Matthew Barth, Prof. Yingbo Hua, Prof. Zak Kassas, Prof. Anastasios Mourikis, Prof. Christian Shelton. Without their help and advice, I cannot have been here.

I want to thank the support (no matter financially or morally) and encouragement from my parents. During the past five years, they are always on my side and help me face every challenge in my life, just like what they had done before I came to US. As the single child of my family, I owe them too much as a son.

I want to thank all the professors and teaching assistants in Mathematics Department who have taught me. Without the rigorous training in mathematics, I cannot complete the part of my dissertation which I am most proud of.

I want to thank the major collaborators in my graduate study: Dr. Wenjie Dong, Dr. Yuanyuan Zhao, Mr. Paul Miller, Dr. Dongfang Zheng, Dr. Arvind Ramanandan, Dr. Anning Chen. I especially thank Dr. Sheng Zhao, without whose help on

To my parents.

ABSTRACT OF THE DISSERTATION

A New Approach for Computationally Efficient and Reliable Carrier Integer
Ambiguity Resolution in GPS/INS

by

Yiming Chen

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, December 2014
Professor Jay Farrell, Chairperson

This dissertation considers reliable real-time navigation with Real-Time Kinematic
(RTK) GPS aided Inertial Navigation System (INS). To improve the accuracy and
reliability of the aided INS, a Contemplative Real-Time (CRT) framework is pro-
posed by combining the conventional Kalman Filtering and the Bayesian smoothing
which considers all the navigation information over a time window. To facilitate the
formulation and the solution of the CRT problem, a probabilistic graphical model
called Factor Graph is utilized. To enhance the robustness of the navigation system
to faulty measurements, a novel robust graph optimization method referred as Hy-
pothesis Test aided Least Soft-thresholding Square (HT-LSS) is proposed. Due to
the integer ambiguity inherent in the GPS carrier phase measurements, in this work
the classical Factor Graph modeling is extended for RTK GPS/INS applications by
incorporating integer unknowns. Nonlinear Mixed Integer Least Square (NMILS)

is required to solve the CRT RTK GPS/INS problem. The major contribution of this thesis is the proposition of a novel Common-Position-Shift method to reduce the computational cost of the NMILS in this problem. In addition, a robust real-time differential correction computation approach was developed for reliable DGPS/RTK applications with internet transported differential information in RTCM and Ntrip standard. The proposed algorithms are evaluated using data acquired with the sensor platform mounted on an automotive vehicle to illustrate the performance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

On September 25, 2012, California Governor Jerry Brown signed legislation that allows self-driving cars (a.k.a. autonomous car, driverless car) to operate on the states roads [8], joining Nevada and Florida, to become the third US state that permits vehicles to drive themselves on public roads [61]. This bill legalizes and encourages the research, development, on-road tests and future volume production of the technologies that enable autonomous vehicles in California. These same technologies enable additional semi-autonomous surface vehicle applications that will enhance safety and mobility while decreasing environmental impact: signal phase and timing, collision avoidance, parking location, cooperative driving, etc. One of the most critical required technologies for autonomous vehicles is accurate and reliable estimation of vehicular navigation state with position accuracy to at the decimeter level (i.e., where-in-lane). Improvements in the accuracy and reliability of land vehicle navigation could not on-

ly accelerate industry innovation, stimulate the local economy but also benefit daily traffic and transportation of Californians significantly in terms of safety and efficiency. Furthermore, the navigation and autonomous control techniques developed for self-driving cars can be extended to other robotics research and development, e.g., agriculture automation and unmanned aerial vehicles.

To achieve reliable land vehicle navigation, various sensors can be integrated on board: Global Navigation Satellite System (GNSS) receiver, odometer, radar, LIDAR, inertial sensors (IMU), Doppler velocity log (DVL) etc. Among these, Global Positioning Systems aided Inertial Navigation Systems (GPS/INS) has proven an efficient integration for navigation purpose, due to their complementary natures [22]. The global positioning accuracy of GPS/INS depends on the performance of the GPS subsystem. A well-designed GPS receiver typically can reach 3-8 meters positioning accuracy with the the U.S. Global Positioning System (GPS) Standard Positioning Service (SPS) [3]. To achieve reliable high precision positioning, differential GPS (DGPS) is a promising technique. With a base station in the range of a few tens of kilometers, DGPS accuracy is in the order of $1m$ (1 σ), growing at the rate of $1m$ per $150km$ of separation [36]. The user can either set up a proprietary base station or utilize the publicly available differential information sources, like Continuously Operating Reference Station (CORS) [60] and Nationwide Differential Global Positioning System (NDGPS) [56]. To facilitate the internet transport of the differential information, international standards e.g. RTCM [59] and Ntrip [57] were proposed

and published. Although the accessibility of internet has been enhance significantly through the vast development of the network infrastructure and the mobile communication, delays of the differential information still impairs the real-time performance of the Differential GPS. Besides, the frequency of measurements broadcast from base station may be lower than that of the rover receiver, such that part of rover measurements cannot be compensated by the base correction. Instead of using Geodetic grade equipment to establish base station, deploying low cost base station to realize DGPS or even Real-Time Kinematic (RTK) positioning is also under considerable effort in industry [63]. When low cost antenna and receiver is used and deployed in unstable electromagnetic environments, base measurements may contains outliers which will also impair the DGPS performance. Thus, robust DGPS correction approach for real-time application is highly required.

Compared with the measurements broadcast from the base stations which are usually established under noise-less environment (e.g. top of mountains and high buildings), the measurements from the rover receiver are more prone to contain faulty data, especially when the rover is under highly dynamic motion or in unideal electromagnetic environment like urban canyon or under heavy foliage. Thus, outlier handling technique is also highly required for GPS/INS to ensure reliability and robustness of the system.

With GPS receivers which can provide carrier phase measurements, Real-Time Kinematic (RTK) technique is able to achieve centimeter accuracy in realtime [65].

However, the RTK technique requires fast and precise on-the-fly integer ambiguity resolution which figures out the unknown cycles numbers in carrier phase measurements. Once the integer ambiguity resolved, the carrier phase measurements work just as the pseudorange but provide much more precise range information (centimeters versus submeters or meters). The integer ambiguity has to be resolved by involving an integer least square (ILS) minimization. A standard ILS problem is NP-hard and then the resolution could be extremely time consuming. In addition, reliable integer ambiguity resolution requires enough inlier satellite measurements. For RTK applications with dual-frequency receivers, the integers can be resolved efficiently in real-time with the aids of the wide-lane phase measurements and applying standard Integer Least Square (ILS) approach, e.g. LAMBDA [33] and MLAMBDA [9]. However, for the single-frequency case, the integer ambiguity becomes pretty challenging depending on only the halved number of GPS measurements and only the L1 carrier phase measurements. In [64], the RTK performance of various combinations of low cost antennas and receivers are evaluated. It was stated that a single-frequency receiver requires at least several minutes to achieve time-to-first-fix (TTFF) with ambiguity resolution in the conclusion of [64]. In [75, 76], a closed form expression of the single-frequency Ambiguity Dilution of Precision (ADOP) is presented to measure ambiguity resolution performance empirically. Still, a large number of epochs is reported in [76] to guarantee the ambiguity resolution success rate even in short baseline situation. However, in [76] it is also concluded that multi-GNSS (e.g. GPS+Beidou)

could improve the ambiguity resolution significantly. Along with the existing GPS, GLONASS, QZSS and the coming full constellations Beidou and Galileo, low cost single-frequency multi-GNSS RTK seems promising. Thus, improvements on the accuracy of integer ambiguity resolution would enable more widely application of RTK technique, with the development of the GNSS constellations and receivers.

GPS (or GNSS) can provide global positioning, velocity and timing information, but in a low frequency (oftentimes 1-2 Hz). For real-time motion planning and automatic control of land vehicles, high frequency navigation state estimates are required. So, inertial navigation systems (INS) which based on high frequency inertial measurement unit (IMU) are usually incorporated with GPS for real-time navigation application. Besides the high frequency position and velocity estimates, the attitude estimates of the vehicle can also be provided by INS. To fuse the navigation information from INS and the aiding sources (e.g., GPS), Extended Kalman Filtering is often used as a standard framework for INS [21, 22, 53]. However, it is well known that EKF has limitations, especially those caused by linearizations [18]. One major disadvantage is that that EKF cannot correct previous wrong linearization points, and thus performs poorly or even collapse when strong nonlinearities exist. Thus, to improve the accuracy and reliability of general INS and then that of GPS/INS, better sensor fusion framework is required.

Considering the above challenges within current GPS/INS research and development, the major contribution of this dissertation is the proposition of a novel

Common-Position-Shift (CPS) approach for solving integer ambiguity over a time window reliably and efficiently. This dissertation also has the following other contributions to improve the accuracy, reliability and robustness of INS, especially for land vehicle applications with GPS aiding:

- Instead of the traditional filtering frame which only considers one epoch aiding measurements, Bayesian smoothing with all navigation information over a time window is applied for INS. Then, for RTK GPS/INS case, the integer ambiguity resolution is executed with a window of IMU and GPS measurements which provide more constraints and improve the accuracy and reliability.

- Probabilistic graphical modeling is introduced to represent the more complicated estimation problem in the smoothing for aided INS. Furthermore, Factor Graph modeling which is widely applied in robotics community, is extended for the RTK GPS/INS application by incorporating the integral unknowns to resolve integer ambiguity.

- To overcome the linearization issue inherited from EKF, iterative nonlinear optimization is applied to solve the *Maximum-a-Posteriori* estimation derived from the graphical model. Thus, previous wrong linearization points can corrected when more measurements included in the optimization of the later windows. For RTK GPS/INS case, the method of solving the Nonlinear Mixed Integer Least Square (NMILS) for smoothing is proposed.

- A novel robust optimization approach is proposed by introducing hypothesis tests to improve the existing methods. This method can be applied in graph optimization base INS, and also fit general least square problems.

- A Contemplative Real-Time (CRT) framework, which combines traditional filtering and the robust graph optimization based smoothing, is proposed for general INS and implemented especially for RTK GPS/INS.

- A novel base correction computation approach is proposed for real-time application with enhanced robustness.

The organization of this dissertation is as follows. Chapter 2 proposes the novel approach for robust base correction computation ensuring real-time performance. Chapter 3 presents the graphical modeling and optimization approach for general Inertial Navigation Systems with aiding measurements. Chapter 4 proposes a hypothesis test based robust optimization approach and uses graph optimization based GPS/INS as an implementation example. Chapter 5 presents the Contemplative Real-Time framework for INS and especially for RTK GPS/INS. Chapter 6 presents a Common-Position-Shift method which could reduce the computational cost of integer ambiguity resolutions in Chapter 5. Chapter 7 concludes this dissertation and discusses potential future works.

Chapter 2

Internet Transport based

Differential GPS

As the most widely applied Global Navigation Satellite System (GNSS), Global Positioning System (GPS) plays crucial roles in modern society by providing global Positioning-Velocity-Timing (PVT) services. A well-designed GPS receiver with antenna typically can reach 3-8 meter positioning accuracy with the the U.S. Global Positioning System (GPS) Standard Positioning Service (SPS) [3].

To achieve reliable higher precision positioning, differential GPS (DGPS) is a promising. The DGPS relies on extra infrastructure which often called base station. A base station is another set of GPS receiver and antenna but it is stationary during the rover operation. The DGPS techniques aim to remove the spatial-common errors between the rover receiver and the base receiver (see Section 2.2). Standard DGPS

requires the global position of the base station is well-surveyed such that the spatial-common errors can be calculated precisely. With a base station in the range of a few tens of kilometers, DGPS accuracy is in the order of $1m$ ($1\ \sigma$), growing at the rate of $1m$ per $150km$ of separation [36].

In practice, the user can either set up a base station on their own [21, 78] or utilize the publicly available correction service, e.g. Continuously Operating Reference Station (CORS) [60], Nationwide Differential Global Positioning System (NDGPS) [56] and EUREF [19]. To facilitate the internet transport of the differential information, international standards e.g. RTCM [59] and Ntrip [57] were proposed and published. As the mobile communication networks (4G or WiFi) becomes readily available, the DGPS technique can be used in various scenarios.

This chapter first reviews the standard differential GPS technique and measurement models in Section 2.1. Then, the internet transport standard for differential GPS information is introduced. Finally, a novel differential correction computation method is proposed to utilize the internet transported differential information in real-time application efficiently and robustly.

2.1 GPS Measurement Models

GPS measurements are made through estimating the travel time of the electro-magnetic signals broadcast from the satellite vehicle antenna to the rover antenna.

In this dissertation, two major types of GPS measurements are used: pseudorange measurements and carrier phase measurements.

2.1.1 Pseudorange Measurements

The L1 and L2 pseudorange measurements for the i-th satellite at time t can be modeled as

$$\tilde{\rho}_{r_1}^i(t) = \|\boldsymbol{p}_r(t) - \boldsymbol{p}^i(t)\|_2 + c\delta t_r(t) + \frac{f_2}{f_1} I_r^i(t) + T_r^i(t) + M_{\rho_1}^i(t) + n_{\rho_1}^i(t), \quad (2.1)$$

$$\tilde{\rho}_{r_2}^i(t) = \|\boldsymbol{p}_r(t) - \boldsymbol{p}^i(t)\|_2 + c\delta t_r(t) + \frac{f_1}{f_2} I_r^i(t) + T_r^i(t) + M_{\rho_2}^i(t) + n_{\rho_2}^i(t), \quad (2.2)$$

where

- $\|\boldsymbol{p}_r - \boldsymbol{p}^i\|_2$ is the geometric distance between the rover position $\boldsymbol{p}_r \in \mathbb{R}^3$ and the i-th satellite vehicle position $\boldsymbol{p}^i \in \mathbb{R}^3$,

- c is the light speed and $\delta t_r \in \mathbb{R}$ is the receiver clock bias which is identical to all channels of the receiver,

- $f_1 = 1575.42 MHz$ and $f_2 = 1227.60 MHz$ are the frequencies for L1 and L2 carrier for GPS signal,

- I_r^i is the Ionospheric delay caused by the ionization effect in the Ionosphere,

- T_r^i is the Tropospheric delay caused by the block from the Troposphere for the L-Band signal,

- $M_{\rho_1}^i$ and $M_{\rho_2}^i$ are multipath errors caused by signal reflections,

- $n_{\rho_1}^i$, $n_{\rho_2}^i \sim \mathcal{N}(0, \sigma_\rho^{i\,2})$ is the (non-common mode) measurement noise.

In practice, the position of the satellite vehicle is estimated from the orbital data which is broadcasted in a data format called ephemeris. Based on the ephemeris data, the satellite orbit can be fitted through the Kepler model [3]. The fitted error in each satellite orbit is unavoidable and called ephemeris error E^i. Also, the clock on the satellite is also estimated with error. Thus, the true range between the satellite vehicle and the rover can be represented as

$$\|\boldsymbol{p}_r - \boldsymbol{p}^i\|_2 = \|\boldsymbol{p}_r - \hat{\boldsymbol{p}}^i\|_2 + E^i + c\delta t^i, \tag{2.3}$$

where $\hat{\boldsymbol{p}}^i$ is the estimated satellite vehicle position from the ephemeris and δt^i is the satellite clock error.

The various errors in eqn. (2.1 - 2.1) can be divided into two categories: common mode and non-common mode. The satellite related errors E^i and $c\delta t^i$ are common to all receivers using the same ephemeris. The atmospheric errors I_r^i and T_r^i are common to spatial nearby ($< 15\text{-}20km$) receivers. The multipath errors $M_{\rho_j}^i$ and receiver noise $n_{\rho_j}^i$ depending on the local electromagnetic environment are non-common between different receivers. Rewriting the common mode error for the i-th satellite as

$$E_{cm1}^i \triangleq E^i + c\delta t^i + \frac{f_2}{f_1}I_r^i + T_r^i, \tag{2.4}$$

$$E_{cm2}^i \triangleq E^i + c\delta t^i + \frac{f_1}{f_2}I_r^i + T_r^i, \tag{2.5}$$

the eqn. (2.1) and (2.2) can be simplified as

$$\tilde{\rho}^i_{r_1}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + E^i_{cm1}(t) + M^i_{\rho_1}(t) + n^i_{\rho_1}(t), \qquad (2.6)$$

$$\tilde{\rho}^i_{r_2}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + E^i_{cm2}(t) + M^i_{\rho_2}(t) + n^i_{\rho_2}(t). \qquad (2.7)$$

Table 2.1.1 from [22] indicates the magnitude of the above errors.

2.1.2 Carrier Phase Measurements

The L1 and L2 carrier phase measurements $\tilde{\varphi}^i_{r1}$ and $\tilde{\varphi}^i_{r2}$ for the i-th satellite at time t can be modeled as

$$\lambda_1 \tilde{\varphi}^i_{r1}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + \lambda_1 N^i_1(t) + E^i_{cm3}(t) + M^i_{\varphi_1}(t) + n^i_{\varphi_1}(t), \quad (2.8)$$

$$\lambda_2 \tilde{\varphi}^i_{r2}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + \lambda_2 N^i_2(t) + E^i_{cm4}(t) + M^i_{\varphi_2}(t) + n^i_{\varphi_2}(t), \quad (2.9)$$

where

- λ_1 and λ_2 are the wavelength of the corresponding carrier signals,

- N^i is the ambiguous integers representing the unknown number of whole cycles,

Common Mode Errors	L1, σ, meters
Ionosphere	7-10
Troposphere	1
Sv Clock	2
Sv Ephemeris	2
Non-common Mode Errors	
Multipath	0.1-3.0
Receiver Noise	0.1-0.7

Table 2.1: Table of User Range Error (URE) standard deviation.

- E_{cm3}^i and E_{cm4}^i are common mode errors similar to E_{cm1}^i and E_{cm2}^i detailed in Section 2.1.1

$$E_{cm3}^i = E^i + c\delta t^i - \frac{f_2}{f_1}I_r^i + T_r^i, \qquad (2.10)$$

$$E_{cm4}^i = E^i + c\delta t^i - \frac{f_1}{f_2}I_r^i + T_r^i, \qquad (2.11)$$

- $M_{\varphi_1}^i$, $n_{\varphi_1}^i$, $M_{\varphi_2}^i$, $n_{\varphi_2}^i$ are non-common mode errors similar to those of pseudorange measurements.

Note that the magnitude of $M_{\varphi_1}^i$, $n_{\varphi_1}^i$, $M_{\varphi_2}^i$, $n_{\varphi_2}^i$ are typically less of 1% of the respective errors in pseudorange measurements [22]. Thus, carrier phase measurements have much lower noise level but are biased by the unknown integer ambiguity $\{N^i\}$. One crucial fact about the integer N^i is that when the Phase-Lock-Loop (PLL) of the corresponding channel for the $i-$th satellite is maintained (i.e., no cycle-slip happens) in the receiver, $N^i(t) \equiv N^i$ is constant. So, if N^i is estimated in previous epoch and no cycle-slip happens, this estimated integer should be used for current and later epochs. In operation, GPS receiver keeps reporting the lock status of the PLL.

2.2 DGPS and Networked Transport Standards

This section reviews the differential GPS technique and the Internet transportation standard and framework for GNSS differential information transportation.

2.2.1 Differential GPS

In DGPS, it is assumed that there exists a nearby (within 15-20km) sationary (i.e., $\boldsymbol{p}_b(t) \equiv \boldsymbol{p}_b \in \mathbb{R}^3$) base station which can provide GPS measurements from the base station receiver to the rover. In other words, the following measurements are available to the rover,

$$\tilde{\rho}_{b_1}^i(t) = \|\boldsymbol{p}_b - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_b(t) + E_{cm1}^{ib}(t) + M_{\rho_1}^{ib}(t) + n_{\rho_1}^{ib}(t), \tag{2.12}$$

$$\tilde{\rho}_{b_2}^i(t) = \|\boldsymbol{p}_b - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_b(t) + E_{cm2}^{ib}(t) + M_{\rho_2}^{ib}(t) + n_{\rho_2}^{ib}(t), \tag{2.13}$$

$$\lambda_1\tilde{\varphi}_{b1}^i(t) = \|\boldsymbol{p}_b - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_b(t) + \lambda N_1^{ib}(t) + E_{cm3}^{ib}(t) + M_{\varphi_1}^{ib}(t) + n_{\varphi_1}^{ib}(t), \tag{2.14}$$

$$\lambda_2\tilde{\varphi}_{b2}^i(t) = \|\boldsymbol{p}_b - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_b(t) + \lambda N_2^{ib}(t) + E_{cm4}^{ib}(t) + M_{\varphi_2}^{ib}(t) + n_{\varphi_2}^{ib}(t). \tag{2.15}$$

Base stations should be established in a good electromagnetic environment (e.g. on the top of the hill or high building with open sky, no multipath effects), such that it is valid to assume that $M_{\rho_1}^{ib} = M_{\rho_2}^{ib} = M_{\varphi_1}^{ib} = M_{\varphi_2}^{ib} \equiv 0$. Furthermore, the base position $\boldsymbol{p}_b$ should be well surveyed with respect to the global frame, such that with the differential technique, more precise global positioning can be realized. Herein, with the known base station position $\boldsymbol{p}_b$, four corrections are defined for later convenience,

$$\alpha_1^i(t) \triangleq \tilde{\rho}_{b_1}^i(t) - \|\boldsymbol{p}_b - \hat{\boldsymbol{p}}^i(t)\|_2, \tag{2.16}$$

$$\alpha_2^i(t) \triangleq \tilde{\rho}_{b_2}^i(t) - \|\boldsymbol{p}_b - \hat{\boldsymbol{p}}^i(t)\|_2, \tag{2.17}$$

$$\alpha_3^i(t) \triangleq \lambda_1\tilde{\varphi}_{b1}^i(t) - \|\boldsymbol{p}_b - \hat{\boldsymbol{p}}^i(t)\|_2, \tag{2.18}$$

$$\alpha_4^i(t) \triangleq \lambda_2\tilde{\varphi}_{b2}^i(t) - \|\boldsymbol{p}_b - \hat{\boldsymbol{p}}^i(t)\|_2. \tag{2.19}$$

Then, with unsubstantial multipath effects around the base station antenna, these four corrections can be modeled as

$$\alpha_1^i(t) = E_{cm1}^{ib}(t) + c\delta t_b(t) + n_{\rho_1}^{ib}(t), \tag{2.20}$$

$$\alpha_2^i(t) = E_{cm2}^{ib}(t) + c\delta t_b(t) + n_{\rho_2}^{ib}(t), \tag{2.21}$$

$$\alpha_3^i(t) = E_{cm3}^{ib}(t) + c\delta t_b(t) + \lambda N_1^{ib}(t) + n_{\varphi_1}^{ib}(t), \tag{2.22}$$

$$\alpha_4^i(t) = E_{cm4}^{ib}(t) + c\delta t_b(t) + \lambda N_2^{ib}(t) + n_{\varphi_2}^{ib}(t). \tag{2.23}$$

If the base station is close to the rover, the following identity is valid

$$E_{cm1}^i(t) = E_{cm1}^{ib}(t), \ E_{cm2}^i(t) = E_{cm2}^{ib}(t), \ E_{cm3}^i(t) = E_{cm3}^{ib}(t), \ E_{cm4}^i(t) = E_{cm4}^{ib}(t), \tag{2.24}$$

since the satellite related errors E^i and $c\delta t^i$ are common to all receivers at the same epoch, and the atmospheric errors I_r^i and T_r^i depend on the locations. If $\{\alpha_j^i(t)\}_{j=1}^4$ is well known, then the raw rover measurements at the same epoch t can be compensated as,

$$^\alpha\tilde{\rho}_{r_1}^{id} = \tilde{\rho}_{r_1}^i - \alpha_1^i, \ ^\alpha\tilde{\rho}_{r_2}^{id} = \tilde{\rho}_{r_2}^i - \alpha_2^i, \ \lambda_1{}^\alpha\tilde{\varphi}_{r1}^{id} = \lambda_1\tilde{\varphi}_{r1}^i - \alpha_3^i, \ \lambda_2{}^\alpha\tilde{\varphi}_{r2}^{id} = \lambda_1\tilde{\varphi}_{r2}^i - \alpha_4^i \tag{2.25}$$

The measurements defined in eqn. (2.25) are often referred as the *single-differenced* measurements with the model

$$^{\alpha}\tilde{\rho}_{r_1}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_{rb}(t) + M_{\rho_1}^i(t) + n_{\rho_1}^{id}(t), \qquad (2.26)$$

$$^{\alpha}\tilde{\rho}_{r_2}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_{rb}(t) + M_{\rho_2}^i(t) + n_{\rho_2}^{id}(t), \qquad (2.27)$$

$$\lambda_1{}^{\alpha}\tilde{\varphi}_{r1}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_{rb}(t) + \lambda_1 N_1^{id}(t) + M_{\varphi_1}^i(t) + n_{\varphi_1}^{id}(t), \qquad (2.28)$$

$$\lambda_1{}^{\alpha}\tilde{\varphi}_{r2}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_{rb}(t) + \lambda_2 N_2^{id}(t) + M_{\varphi_2}^i(t) + n_{\varphi_2}^{id}(t), \qquad (2.29)$$

where $\delta t_{rb} \triangleq \delta t_r - \delta t_b \in \mathbb{R}$ and $N_j^{id} \triangleq N_j^i - N_j^{ib} \in \mathbb{N}, j = 1, 2$ are the differenced clock bias and the differenced integer between the rover receiver and the base receiver, as well as the noise

$$n_{\rho_1}^{id} \triangleq n_{\rho_1}^i - n_{\rho_1}^{ib}, \; n_{\rho_2}^{id} \triangleq n_{\rho_2}^i - n_{\rho_2}^{ib}, \; n_{\varphi_1}^{id} \triangleq n_{\varphi_1}^i - n_{\varphi_1}^{ib}, \; n_{\varphi_2}^{id} \triangleq n_{\varphi_2}^i - n_{\varphi_2}^{ib}.$$

In DGPS applications, the single-differenced measurements are used for positioning practice instead of using the raw measurements in eqn. (2.1, 2.2, 2.8, 2.9). DGPS provides higher accuracy since the common mode errors are compensated. Since the common mode errors are time-varying, the rover measurements should be compensated with the base correction at the same epoch. In post-processing, the time-stamp correspondence can be realized by matching the GPS time-of-week of the measurements. However, in real-time implementations, the differential data transportation from the base station is delayed. In Section 2.3, a real-time differential correction approach is proposed to handle this issue.

2.2.2 RTCM and Ntrip Standard

To transmit the differential GNSS message, inproprietary and efficient protocols are highly required. For this purpose, the Special Committee 104 on DGNSS of the Radio Technical Commission for Maritime Services (RTCM) proposed a standard for the dissemination of the differential information in binary messages [59]. This standard is usually referred as the RTCM standard, and the current version is 3.1.

"Networked Transport of RTCM via Internet Protocol" (Ntrip) stands for an application level protocol streaming GNSS data over the Internet. Ntrip is a generic, stateless protocol based on the Hypertext Transfer Protocol (HTTP/1.1) and the Real Time Streaming Protocol (RTSP). Ntrip has been designed for disseminating differential correction data (e.g. in the formats of RTCM Special Committee 104) or other kinds of GNSS streaming data to stationary or mobile users over the Internet.

Ntrip consists of three system software components: clients, servers and casters. Fig. 2.1 shows the basic structure of a Ntrip system. In a typical Ntrip system, Ntrip servers transmit the RTCM message generated from base station receivers to the Ntrip Caster. One Ntrip Caster can carry multiple Ntrip servers. Ntrip Clients on the rover would request single or multiple Ntrip streams from Ntrip Caster with valid authorizations. By parsing the RTCM message in the Ntrip stream, the differential information (e.g. base receiver measurements or corrections) can be obtained. Fig. 2.2 shows the basic structure of the DGPS positioning routine on the rover: Ntrip

client communicates with the Ntrip caster to get the required RTCM stream, then the binary RTCM messages are parsed by a parser, then the differential correction is calculated (see Section 2.3), and then after compensating the rover measurements with differential corrections, DGPS positioning can be executed.
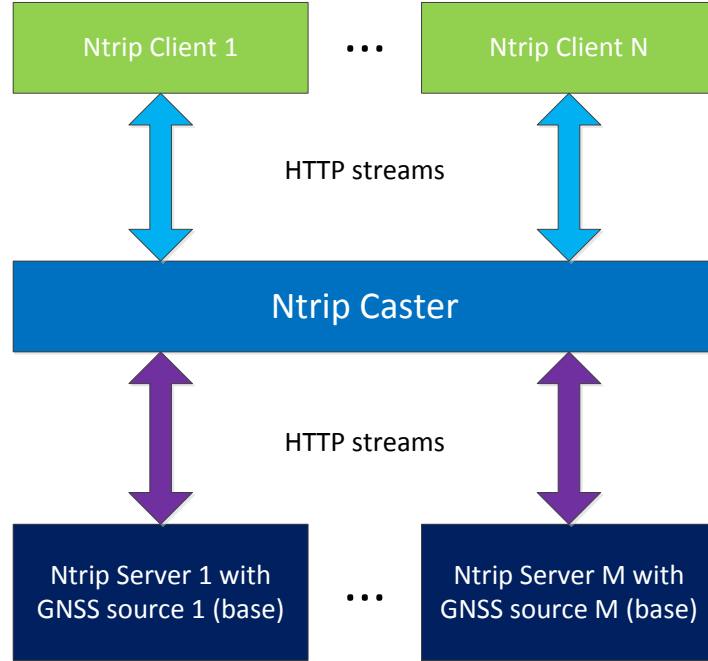
Figure 2.1: Ntrip system structure with Ntrip Casters, Ntrip Servers, Ntrip Clients and base stations.

2.3   Real-Time Robust Correction Calculation

As shown in Section 2.2, the key procedure in DGPS is to compute the correction $\{\alpha_j^i(t)\}_{j=1}^4$. While the access to the internet is quite available in nowadays due to the development of network technology and infrastructures, transport delays of the
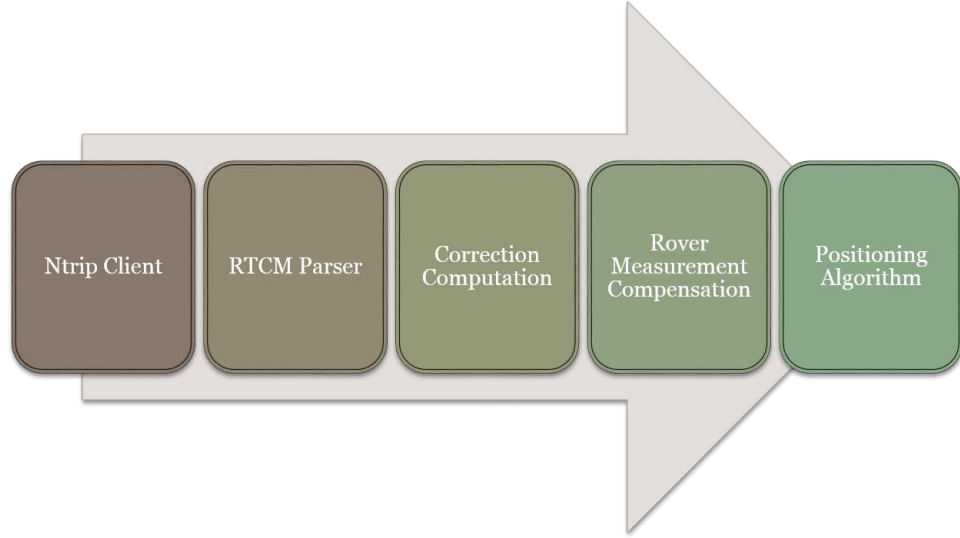
Figure 2.2: Structure of the positioning routine of the rover with a Ntrip client.

differential data still can happen frequently on a fast moving rover, or with limited connections. Thus, the base corrections cannot be calculated in real-time. Furthermore, when the differential data is from a temporary base station equipped with low cost antenna and receiver, or in an noisy electromagnetic environment, there may exist outliers in the base measurements. This section presents a base correction calculation method which can improve the real-time performance of DGPS and ensure robustness of the correction calculation.

Without loss of generality, let us only consider the correction calculation of L1 code and phase measurements. The proposed method can be extended to dual frequency cases straightforwardly. Then, the description of the correction calculation problem can be stated as follows: given that

- the surveyed base ECEF (Earth-Centered-Earth-Fixed) position $\boldsymbol{p}_b \in \mathbb{R}^3$

- circular buffers for each satellite maintain the code and phase measurements $\{\tilde{\rho}^i_{b_1}(t_{k-d})\}, k = 1, \ldots, K$ and $\{\lambda_1 \tilde{\varphi}^i_{b1}(t_{k-d})\}, k = 1, \ldots, K$, where $d \in \mathbb{Z}^+$ indicates the delay in epoch numbers and K is the circular buffer length,

- the satellite vehicle ECEF positions $\{\hat{\boldsymbol{p}}^i(t_{k-d})\}$ estimated with the ephemeris from the rover receiver,

the real-time corrections $\{\alpha^i_j(t_K)\}^4_{j=1}$ defined in eqn. (2.16 - 2.19) are expected to be estimated, where t_K indicates the time-of-week of the latest GPS measurement from the rover receiver.

The above problem statement implies that the real-time differential correction calculation with delayed data is a regression/prediction problem. To decide the proper model for this regression, the patterns of the corrections $\{\alpha^i_j(t_K)\}^4_{j=1}$ are first investigated. In Fig. 2.3, the differential corrections $\alpha^i_1(t)$ for L1 code measurements of different satellites are plotted with different color over a $1000s$ time interval. Two instant observations can be derived from Fig. 2.3:

- $\alpha^i_1(t)$ has a high frequency (fast-varying) and a low frequency (slowly-varying) components;

- the high frequency components for different satellites have similar pattern.

Since the satellite related errors E^i and $c\delta t^i$ and the atmospheric errors I_r^i and T_r^i are all slowly-varying signals, the common mode errors defined in eqn. (2.4 - 2.11) are the low frequency components in $\alpha_1^i(t)$. On the other hand, Fig. 2.4 shows the estimated GPS receiver clock bias in positioning and indicates that the base receiver clock bias is in the high frequency components of $\alpha_1^i(t)$. Due to the mixture of low frequency and high frequency components, it is not straightforward to predict $\alpha_1^i(t)$.

Define

$$\beta_1^i(t) \triangleq \alpha_1^i(t) - c\delta t_b(t) = E_{cm1}^{ib}(t) + n_{\rho_1}^{ib}(t), \tag{2.30}$$

$$\beta_2^i(t) \triangleq \alpha_2^i(t) - c\delta t_b(t) = E_{cm2}^{ib}(t) + n_{\rho_2}^{ib}(t), \tag{2.31}$$

$$\beta_3^i(t) \triangleq \alpha_3^i(t) - c\delta t_b(t) = E_{cm3}^{ib}(t) + \lambda N_1^{ib}(t) + n_{\varphi_1}^{ib}(t), \tag{2.32}$$

$$\beta_4^i(t) \triangleq \alpha_4^i(t) - c\delta t_b(t) = E_{cm4}^{ib}(t) + \lambda N_2^{ib}(t) + n_{\varphi_2}^{ib}(t), \tag{2.33}$$

then another version of single-differenced measurements can be defined as

$$^\beta\tilde{\rho}_{r_1}^{id} = \tilde{\rho}_{r_1}^i - \beta_1^i, \ ^\beta\tilde{\rho}_{r_2}^{id} = \tilde{\rho}_{r_2}^i - \beta_2^i, \ \lambda_1{}^\beta\tilde{\varphi}_{r1}^{id} = \lambda_1\tilde{\varphi}_{r1}^i - \beta_3^i, \ \lambda_2{}^\beta\tilde{\varphi}_{r2}^{id} = \lambda_1\tilde{\varphi}_{r2}^i - \beta_4^i \tag{2.34}$$

with the models

$$^\beta\tilde{\rho}_{r_1}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + M_{\rho_1}^i(t) + n_{\rho_1}^{id}(t), \tag{2.35}$$

$$^\beta\tilde{\rho}_{r_2}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + M_{\rho_2}^i(t) + n_{\rho_2}^{id}(t), \tag{2.36}$$

$$\lambda_1{}^\beta\tilde{\varphi}_{r1}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + \lambda_1 N_1^{id}(t) + M_{\varphi_1}^i(t) + n_{\varphi_1}^{id}(t), \tag{2.37}$$

$$\lambda_1{}^\beta\tilde{\varphi}_{r2}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + \lambda_2 N_2^{id}(t) + M_{\varphi_2}^i(t) + n_{\varphi_2}^{id}(t). \tag{2.38}$$
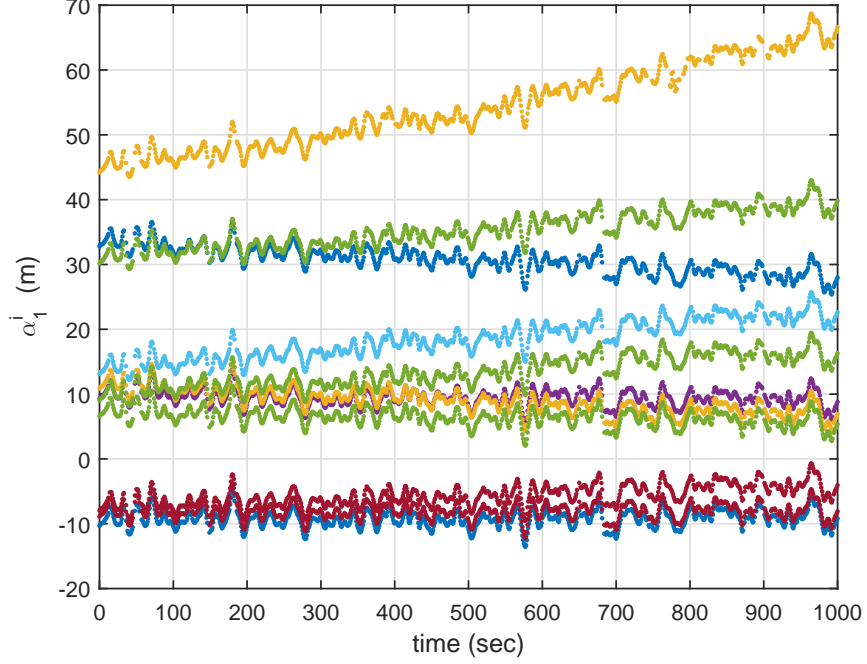
Figure 2.3: All corrections $\{\alpha_1^i(t)\}$ for L1 code measurements over a 1000s interval. Different colors are for different satellites.

Note that the only differences between eqn. (2.26-2.29) and eqn. (2.35-2.38) are that in eqn. (2.35-2.38) the rover receiver biases $c\delta t_r(t)$ are not mixed with the base receiver biases $c\delta t_b(t)$. It can be proved that the positioning results from these two versions of single-differenced measurements are identical. If the base receiver locks are maintained, then $\lambda N_1^{ib}(t) \equiv \lambda N_1^{ib} \in \mathbb{Z}$ and $\lambda N_2^{ib}(t) \equiv \lambda N_2^{ib} \in \mathbb{Z}$ are integral constants. Along with that the common errors $\{E_{cmj}^{ib}(t)\}_{j=1}^{4}$ are slowly-varying variables, a linear regression based method is proposed in the following to estimate the corrections $\{\beta_j^i(t)\}_{j=1}^{4}$ in real-time. Thus, the single-differenced measurements $^{\beta}\tilde{\rho}_{r_1}^{id}$, $^{\beta}\tilde{\rho}_{r_2}^{id}$, $^{\beta}\tilde{\varphi}_{r2}^{id}$, $^{\beta}\tilde{\varphi}_{r2}^{id}$ can be calculated. The procedures of the proposed method are as follows
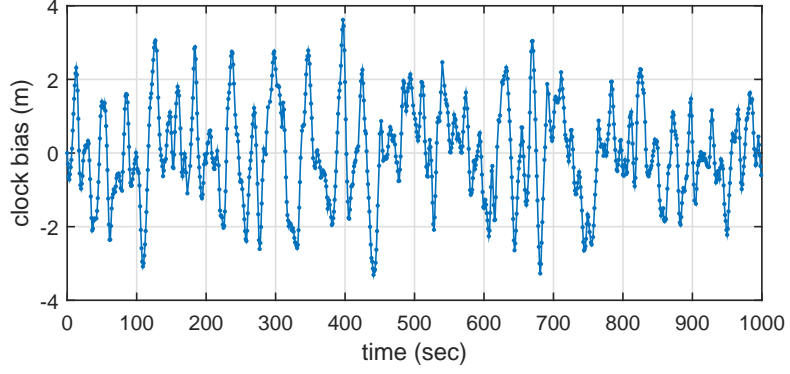
Figure 2.4: GPS clock bias (fine-steering enabled) estimates from single epoch positioning.

- Pass $\{\alpha_1^i(t_{1-d}), \ldots, \alpha_1^i(t_{K-d})\}$ computed through eqn. (2.16) for each satellite to a low-pass filter (e.g. moving average filter) $C(s)$ to obtain the output $\{\hat{\alpha}_1^i(t_{1-d}), \ldots, \hat{\alpha}_1^i(t_{K-d})\}$,

- Take the median of the error $\epsilon_1^i(t) \triangleq \alpha_1^i(t) - \hat{\alpha}_1^i(t)$ over all available satellites $(i = 1, \ldots, m)$ to get an estimate of the base clock bias

$$clk(t) \triangleq median(\epsilon_1^i(t)), \qquad (2.39)$$

and Fig. shows an example for this step.

- Evaluate $\{\gamma_j^i(t)\}_{j=1}^4$ as

$$\gamma_j^i(t) \triangleq \alpha_j^i(t) - clk(t)$$

and keep them in circular buffers with length L. Use $\{\gamma_j^i(t_{K-d-L+1}), \ldots, \gamma_j^i(t_{K-d})\}$ to fit a linear model $\hat{\beta}_j^i(t)$ for estimating $\beta_j^i(t_K)$.

23
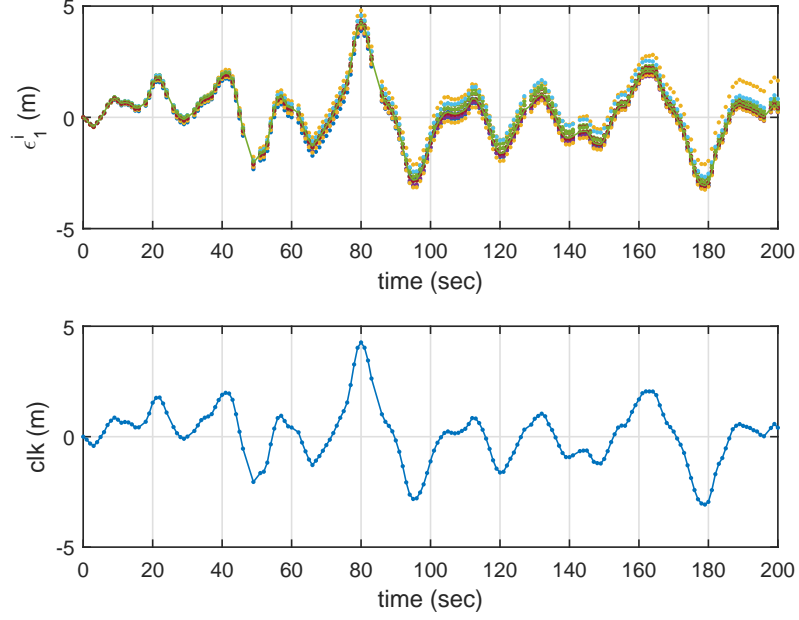
Figure 2.5: Filtered noise $\epsilon_1^i(t)$ and the estimate of the base clock bias $clk(t)$.

- Finally, the rover uses the estimated DGPS measurements

$$^\gamma\tilde{\rho}_{r_1}^{id} = \tilde{\rho}_{r_1}^i - \hat{\beta}_1^i, \ ^\gamma\tilde{\rho}_{r_2}^{id} = \tilde{\rho}_{r_2}^i - \hat{\beta}_2^i, \ \lambda_1{}^\gamma\tilde{\varphi}_{r1}^{id} = \lambda_1\tilde{\varphi}_{r1}^i - \hat{\beta}_3^i, \ \lambda_2{}^\gamma\tilde{\varphi}_{r2}^{id} = \lambda_1\tilde{\varphi}_{r2}^i - \hat{\beta}_4^i$$

(2.40)

with the models

$$^\gamma\tilde{\rho}_{r_1}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + M_{\rho_1}^i(t) + n_{\rho_1}^{id}(t),$$

(2.41)

$$^\gamma\tilde{\rho}_{r_2}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + M_{\rho_2}^i(t) + n_{\rho_2}^{id}(t),$$

(2.42)

$$\lambda_1{}^\gamma\tilde{\varphi}_{r1}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + \lambda_1 N_1^{id} + M_{\varphi_1}^i(t) + n_{\varphi_1}^{id}(t),$$

(2.43)

$$\lambda_1{}^\gamma\tilde{\varphi}_{r2}^{id}(t) = \|\boldsymbol{p}_r(t) - \hat{\boldsymbol{p}}^i(t)\|_2 + c\delta t_r(t) + \lambda_2 N_2^{id} + M_{\varphi_2}^i(t) + n_{\varphi_2}^{id}(t).$$

(2.44)

to do positioning.

24

Note that in eqn. (2.43-2.44), integer ambiguities N_1^{id} and N_2^{id} get rid of the time stamp, since it is assumed that the phase-lock-loop of the base receiver for the i-th satellite is maintained. This assumption is often satisfied since most of base stations are established in a good electromagnetic environment, e.g. on the top of the mountain or high buildings with high quality antenna (see CORS stations [60]). With low cost single frequency receivers in noisy environment as the base, the proposed correction calculation method can enhance the robustness of the differential positioning significantly. The reason is that all data in a window is used for the correction fitting which improves the degree-of-freedom of the estimation problem. As shown in Chapter 4, robust optimization technique can be applied in the proposed correction calculation method to handle faulty measurements caused by unexpected noise or undetected cycle slips.

Fig. 2.6 shows the overall process of the differential correction calculation for L1 code measurements. Comparing $\gamma_1^i(t)$ with $\alpha_1^i(t)$, it can be seen that the high frequency component caused by the clock bias and the measurement noise are removed, such that a linear model can be fitted to calculate $\hat{\beta}_1^i(t)$.

In the rest of this chapter, several example results of differential GPS are presented to illustrate the proposed method. Since only positioning is focused on, in all the following experiments, the rover is stationary over a 30-minute window. The ground truth of rover's global position is surveyed with geodetic grade dual frequency receiver and antenna. The survey data was processed by Online Positioning User Service

(OPUS) [1] aided with measurements from nearby differential stations. The accuracy of the survey is $\pm(0.02, 0.02, 0.02)m$ in North-East-Down directions.

Totally six experiments in no-baseline, short-baseline and long-baseline scenarios were done as shown in the following table. Herein, the baseline is defined as the distance between the rover receiver and the base station receiver. Among the experiments, no-baseline and short-baseline experiment data was collected with UC Riverside Ntrip station (ntrip.engr.ucr.edu@2101) and a rover receiver and antenna on UCR campus. In particular, no-baseline experiments were done by connecting the base and rover receivers to the same antenna. The long-baseline experiment data was collected with ESRI GISA Ntrip station (esricaster.esri.com@2101) and a rover receiver and antenna on UCR campus. Fig. 2.7 and 2.8 show that in no-baseline condition, even there is time delay in Ntrip stream, the base correction computed by the proposed method can cancel the common error and realize unbiased positioning result. Fig. 2.9 and 2.10 show the expected centimeter-level accuracy of the RTK technique. Fig. 2.11 to 2.18 show that although the positioning error of DGPS grows as the baseline expands, the RTK technique with the calculated base corrections still can realize centimeter level positioning accuracy.

In the rest of this dissertation, the proposed base correction computation approach is applied in all GPS/INS implementations.

Table 2.2: Base Correction Computation Experiments

Baseline (km)	Measurement used	Ntrip delay (sec)	Figure number
0	L1 Code	3-5	2.7, 2.8
0	L1 Code+Phase	3-5	2.9, 2.10
< 1	L1 Code	3-5	2.11, 2.12
< 1	L1 Code+Phase	3-5	2.13, 2.14
15	L1 Code	1-2	2.15, 2.16
15	L1 Code+Phase	1-2	2.17, 2.18
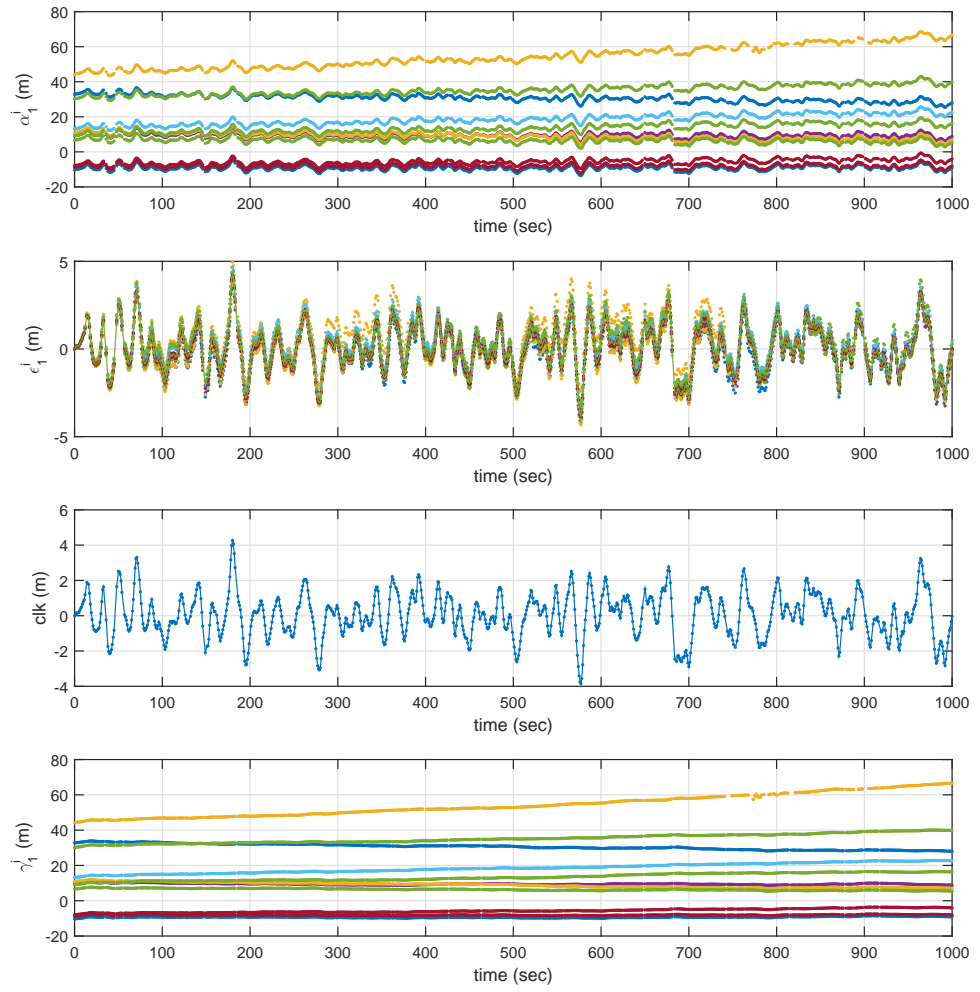
Figure 2.6: Real-time correction computation process.
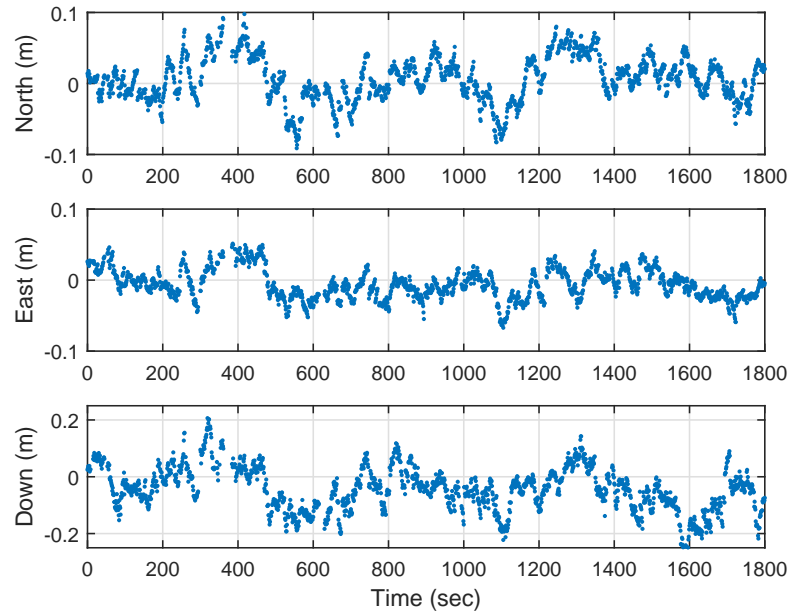
Figure 2.7: L1 DGPS Code-only positioning: no-baseline result.
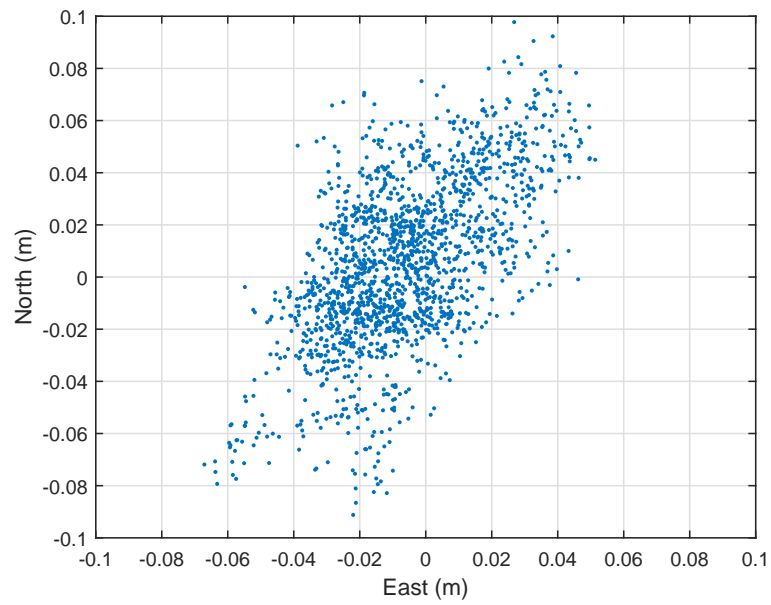
Figure 2.8: L1 DGPS Code-only positioning: no-baseline result on tangent plane.
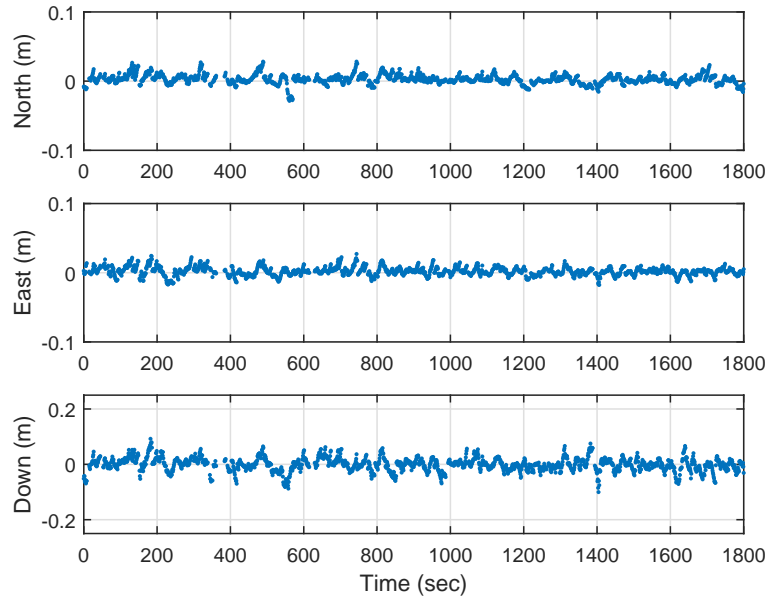
Figure 2.9: L1 DGPS Code+Phase positioning: no-baseline result.
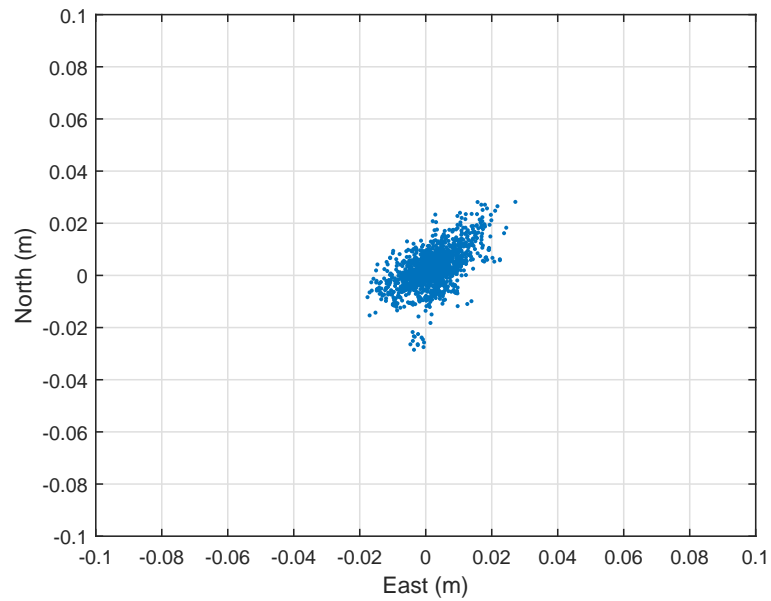
Figure 2.10: L1 DGPS Code+Phase positioning: no-baseline result on tangent plane.
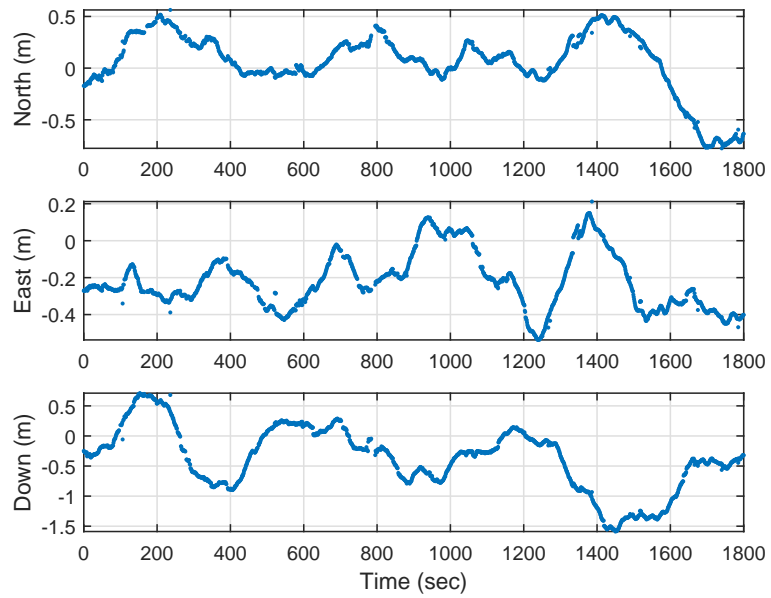
Figure 2.11: L1 DGPS Code-only positioning: short-baseline result.
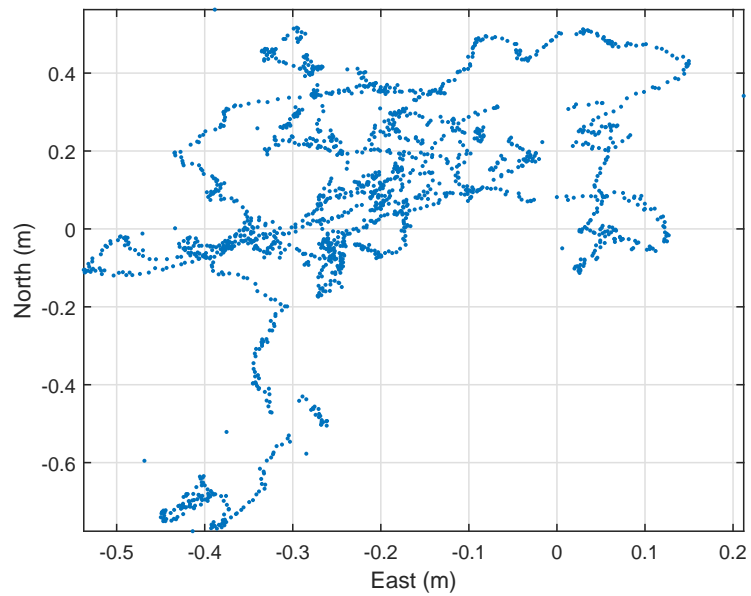
Figure 2.12: L1 DGPS Code-only positioning: short-baseline result on tangent plane.
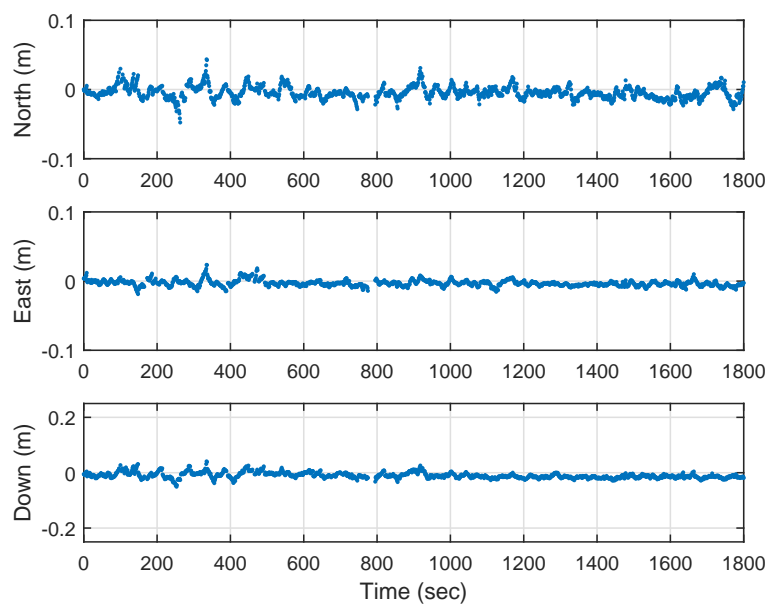
Figure 2.13: L1 DGPS Code+Phase positioning: short-baseline result.
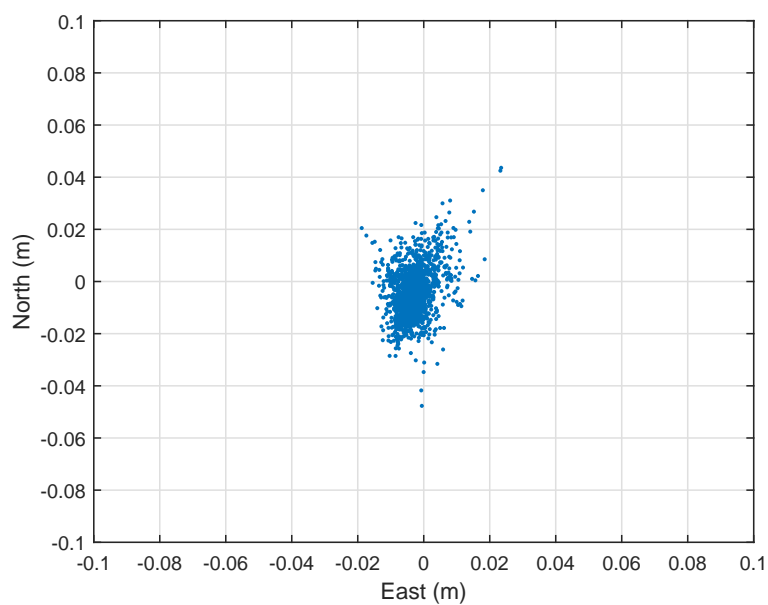
Figure 2.14: L1 DGPS Code+Phase positioning: short-baseline result on tangent plane.
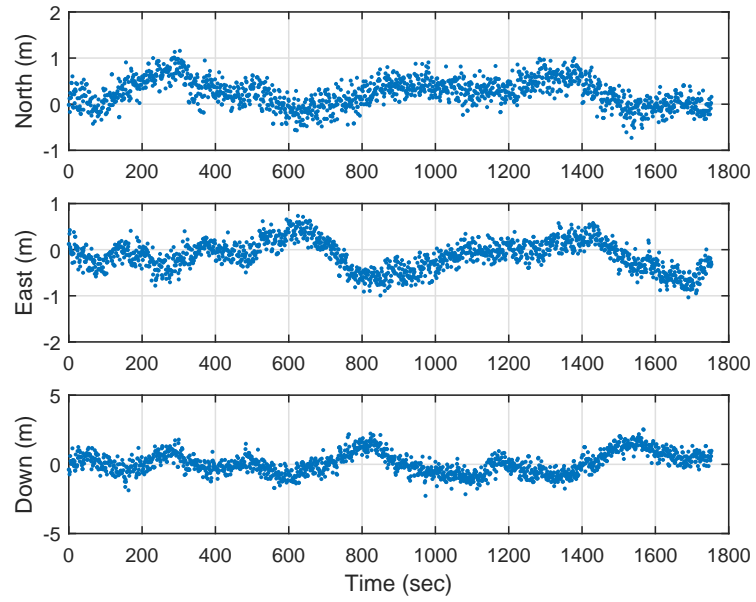
Figure 2.15: L1 DGPS Code-only positioning: long-baseline result.
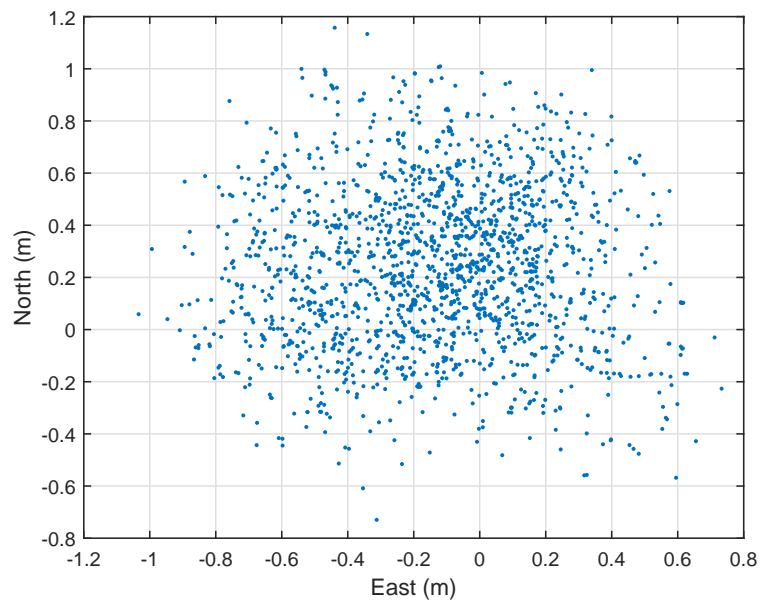
Figure 2.16: L1 DGPS Code-only positioning: long-baseline result on tangent plane.
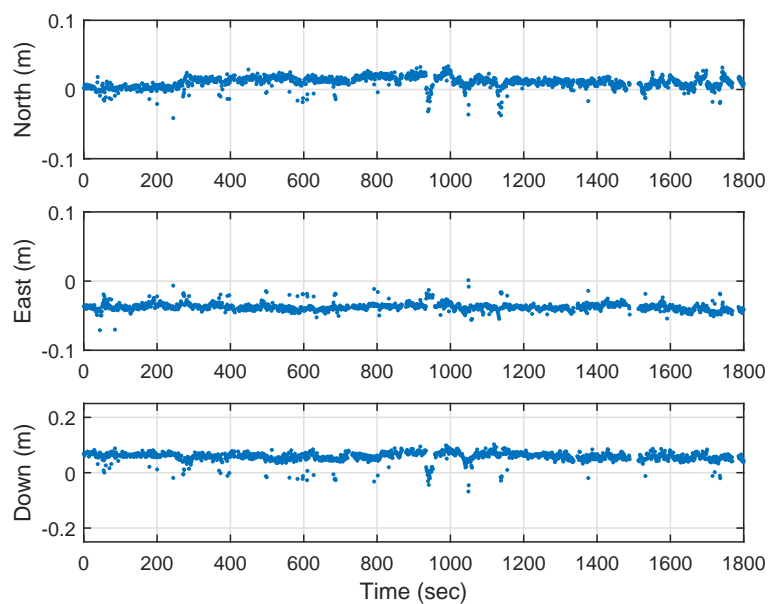
Figure 2.17: L1 DGPS Code+Phase positioning: long-baseline result.
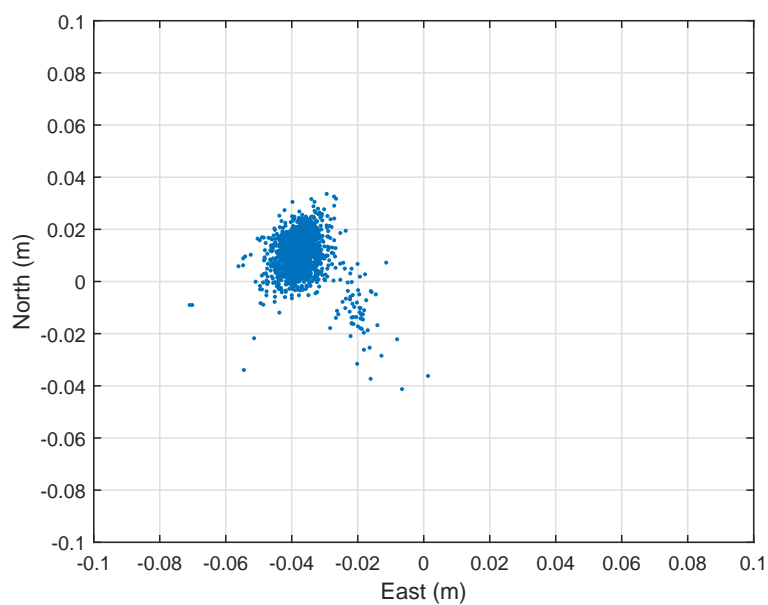
Figure 2.18: L1 DGPS Code+Phase positioning: long-baseline result on tangent plane.

Chapter 3

Graph Optimization based Inertial Navigation Systems (GOINS)

In this chapter, a graph optimization based smoothing method is proposed for Aided Inertial Navigation Systems (INS) [22]. The goal of the smoothing is to improve the accuracy and reliability of the INS which is traditionally constructed under EKF framework. Furthermore, the robust optimization approach proposed in Chapter 4 can applied to enhance the robustness of the system to faulty measurements. To make this dissertation self-contained, this chapter first reviews the EKF based INS mechanism and notation. Then, the graph optimization based INS is proposed and the details of the solution are given.

3.1 Inertial Navigation Systems

Nowadays, Inertial Navigation Systems are widely applied in different areas to provide position, velocity and attitude estimates of robot, land vehicle [21], underwater vehicle[53], and pedestrian [47]. In this dissertation, it is assumed that the INS is based on a 6-axis Inertial Measurement Unit (IMU) which provides high frequency 3D acceleration and angular rate measurements from a 3-axis accelerometer and a 3-axis gyroscope. This section reviews the basic notation and mathematical derivation of EKF based aided INS.

Let $\boldsymbol{x} \in \mathbb{R}^{n_s}$ denote the rover state vector. The kinematic equations for the rover state can be described by

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}), \tag{3.1}$$

where $\boldsymbol{f} : \mathbb{R}^{n_s} \times \mathbb{R}^6 \mapsto \mathbb{R}^{n_s}$ is the kinematic model, $\boldsymbol{u} \in \mathbb{R}^6$ is the vector of accelerations and angular rates (representing the external control input to the system).The function $\boldsymbol{f}$ is accurately known (see Chapter 11 in [22]). Given an initial condition for the state vector $\hat{\boldsymbol{x}}(0) \in \mathbb{R}^n$ and measurements $\tilde{\boldsymbol{u}}$ of $\boldsymbol{u}$, an Inertial Navigation System (INS) propagates an estimate of the vehicle state between aiding measurements as a solution of

$$\dot{\hat{\boldsymbol{x}}} = \boldsymbol{f}(\hat{\boldsymbol{x}}, \tilde{\boldsymbol{u}}), \tag{3.2}$$

where $\hat{\boldsymbol{x}}$ denotes the estimate of $\boldsymbol{x}$. Due to initial condition errors, system calibration errors, and measurement noise, the state estimation error $\delta\boldsymbol{x}(t) = \boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t)$ devel-

ops over time. Although the dynamics and stochastic properties of this estimation error are well understood [22], a brief derivation is presented in this section for the convenience of later discussions and the completeness of this dissertation.

3.1.1 Preliminary

This subsection clarifies some notations and preliminaries [22, 51] may be used in later discussion:

- $\boldsymbol{p}_{a/b} \in \mathbb{R}^3$ denotes the translation from the origin of frame a to that of frame b.

- $\boldsymbol{v}_{a/b} \in \mathbb{R}^3$ denotes the relative velocity of frame b with respect to frame a.

- $\boldsymbol{\omega}_{a/b} \in \mathbb{R}^3$ denotes the relative rate of rotation of frame b with respect to frame a.

- The superscript indicates the frame of reference, e.g., $\boldsymbol{p}_{a/b}^c$ denotes $\boldsymbol{p}_{a/b}$ represented in frame c.

- $\boldsymbol{R}_a^b \in SO(3)^1$ denotes the rotation matrix transforming vectors from frame b to frame a, e.g., $\boldsymbol{p}_{a/b}^b = \boldsymbol{R}_a^b \boldsymbol{p}_{a/b}^a$. The rotation matrix has the following properties: $|\boldsymbol{R}_a^b| = 1$ and $(\boldsymbol{R}_a^b)^\intercal \boldsymbol{R}_a^b = \boldsymbol{I}_3$.

---

[1]$SO(3) = \{\boldsymbol{R} \in \mathbb{R}^{3\times3} | \boldsymbol{R}^\intercal \boldsymbol{R} = \boldsymbol{I}, det(R) = +1\}$ denotes the group of *special orthogonal matrix*, see [51].

- Let $\boldsymbol{\omega} = [\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \boldsymbol{\omega}_3]^\mathsf{T}$, then

$$[\boldsymbol{\omega}\times] \triangleq \begin{bmatrix} 0 & -\boldsymbol{\omega}_3 & \boldsymbol{\omega}_2 \\ \boldsymbol{\omega}_3 & 0 & -\boldsymbol{\omega}_1 \\ -\boldsymbol{\omega}_2 & \boldsymbol{\omega}_1 & 0 \end{bmatrix} \in so(3)^2,$$

  denotes the skew-symmetric matrix associated to $\boldsymbol{\omega}$. A skew-symmetric matrix
  has the property $[\boldsymbol{\omega}\times]^\mathsf{T} = -[\boldsymbol{\omega}\times]$.

- For the angular rate $\boldsymbol{\omega}_{a/b}^b$, it is often defined that $\boldsymbol{\Omega}_{a/b}^b \triangleq [\boldsymbol{\omega}_{a/b}^b\times]$.

- The derivative of the rotation matrix $\boldsymbol{R}_a^b$ has the following relation with the
  angular rate $\boldsymbol{\omega}_{a/b}$ (and then $\boldsymbol{\Omega}_{a/b}$),

$$\dot{\boldsymbol{R}}_a^b(t) = \boldsymbol{R}_a^b(t)\boldsymbol{\Omega}_{a/b}^a = -\boldsymbol{\Omega}_{a/b}^b\boldsymbol{R}_a^b(t). \tag{3.3}$$

- Suppose that the rotation from frame a to frame b is an infinitesimal rotation
  denoted by $\delta\boldsymbol{\theta} = [\delta\theta_1, \delta\theta_2, \delta\theta_3]^\mathsf{T}$, then the rotation matrix from frame a to frame
  b can be approximated as

$$\boldsymbol{R}_a^b = \boldsymbol{I}_3 - \delta\boldsymbol{\Theta}, \tag{3.4}$$

  where $\delta\boldsymbol{\Theta} \triangleq [\delta\boldsymbol{\theta}\times]$.

- Let $\theta \in [-\pi, \pi]$ be an angle, in this dissertation the following notations are used
  for simplicity: $s\theta \triangleq \sin(\theta)$, $c\theta \triangleq \cos\theta$ and $t\theta \triangleq \tan\theta$.

---

[2] $so(3) = \{[\boldsymbol{\omega}\times] \in \mathbb{R}^{3\times3} | \boldsymbol{\omega} \in \mathbb{R}^3\}$ denotes the space of all skew-symmetric matrices, see [51].

3.1.2 Kinematic Model Derivation

The navigation model used in this dissertation is based on the well-understood kinematic relationships for rigid body motion [22]. The origin of the inertial frame is picked coincident with that of the Earth Center Earth Fixed (ECEF) frame, i.e. the translation $\boldsymbol{p}_{i/e} = \boldsymbol{0}$. Suppose that the rover operates on a fixed tangent frame with a North-East-Down (NED) coordinate system, and then the translation $\boldsymbol{p}_{e/t}$ and the rotation $\boldsymbol{R}_{e/t}$ are well-known constants. Then, the rover position and velocity represented in the inertial frame of reference is

$$\boldsymbol{p}_{i/b}^i = \boldsymbol{p}_{i/t}^i + \boldsymbol{R}_t^i \boldsymbol{p}_{t/b}^t, \tag{3.5}$$

$$\boldsymbol{v}_{i/b}^i = \dot{\boldsymbol{p}}_{i/t}^i + \boldsymbol{R}_t^i \left[\boldsymbol{\Omega}_{i/t}^t \boldsymbol{p}_{t/p}^t + \dot{\boldsymbol{p}}_{t/p}^t \right], \tag{3.6}$$

where $\boldsymbol{\Omega}_{i/t}^t = [\boldsymbol{\omega}_{i/t}^t \times]$. Eqn. (3.6) is from eqn. (3.3). In eqn. (3.6), $\boldsymbol{\omega}_{i/t}^t$ represents the rotation rate of the tangent frame t with respect to inertial frame i, in the tangent frame. Then, the acceleration of the rover relative to the inertial frame can be derived by differentiating eqn. (3.6),

$$\dot{\boldsymbol{v}}_{i/b}^i = \ddot{\boldsymbol{p}}_{i/t}^i + \boldsymbol{R}_t^i \boldsymbol{\Omega}_{i/t}^t \left[\boldsymbol{\Omega}_{i/t}^t \boldsymbol{p}_{t/b}^t + \dot{\boldsymbol{p}}_{t/b}^t \right] + \boldsymbol{R}_t^i \left[\dot{\boldsymbol{\Omega}}_{i/t}^t \boldsymbol{p}_{t/b}^t + \boldsymbol{\Omega}_{i/t}^t \dot{\boldsymbol{p}}_{t/b}^t + \ddot{\boldsymbol{p}}_{t/b}^t \right],$$

and it follows that [3]

$$\dot{\boldsymbol{v}}_{i/b}^t = \boldsymbol{R}_i^t \dot{\boldsymbol{v}}_{i/b}^i = \boldsymbol{\Omega}_{i/t}^t \left[\boldsymbol{\Omega}_{i/t}^t \boldsymbol{p}_{t/b}^t + \dot{\boldsymbol{p}}_{t/b}^t \right] + \dot{\boldsymbol{\Omega}}_{i/t}^t \boldsymbol{p}_{t/b}^t + \boldsymbol{\Omega}_{i/t}^t \dot{\boldsymbol{p}}_{t/b}^t + \ddot{\boldsymbol{p}}_{t/b}^t. \tag{3.7}$$

---

[3] $\ddot{\boldsymbol{p}}_{i/t}^i \equiv \boldsymbol{0}$, since the tangent frame is fixed.

With eqn. (3.7), the acceleration of the rover relative to the tangent frame is

$$\dot{\boldsymbol{v}}_{t/b}^t = \ddot{\boldsymbol{p}}_{t/b}^t = \dot{\boldsymbol{v}}_{i/b}^t - \boldsymbol{\Omega}_{i/t}^t \boldsymbol{\Omega}_{i/t}^t \boldsymbol{p}_{t/b}^t - 2\boldsymbol{\Omega}_{i/t}^t \dot{\boldsymbol{p}}_{t/b}^t - \dot{\boldsymbol{\Omega}}_{i/t}^t \boldsymbol{p}_{t/b}^t.$$

Using the definition of the specific force [22]

$$\mathbf{f}_{i/b} = \dot{\boldsymbol{v}}_{i/b} - \mathbf{G}_{i/b}, \tag{3.8}$$

it follows that [4]

$$\begin{aligned} \dot{\boldsymbol{v}}_{t/b}^t &= \mathbf{f}_{i/b}^t + \mathbf{G}_{i/b}^t - \boldsymbol{\Omega}_{i/t}^t \boldsymbol{\Omega}_{i/t}^t \boldsymbol{p}_{t/b}^t - 2\boldsymbol{\Omega}_{i/t}^t \dot{\boldsymbol{p}}_{t/b}^t, \\ &= \boldsymbol{R}_b^t \mathbf{f}_{i/b}^b + \mathbf{g}_{i/b}^t - 2\boldsymbol{\Omega}_{i/t}^t \dot{\boldsymbol{p}}_{t/b}^t, \end{aligned} \tag{3.9}$$

where $\mathbf{g}_{i/b}^t = \mathbf{G}_{i/b}^t - \boldsymbol{\Omega}_{i/t}^t \boldsymbol{\Omega}_{i/t}^t \boldsymbol{p}_{t/b}^t$ is the local gravity vector. In particular, the local gravity vector $\mathbf{g}_{i/b}^t$ is composed of the gravitational force and the centripetal acceleration caused by Earth rotation. The centripetal acceleration $-\boldsymbol{\Omega}_{i/t}^t \boldsymbol{\Omega}_{i/t}^t \boldsymbol{p}_{t/b}^t$ depends on the global location of the rover, since $\boldsymbol{\omega}_{i/t}^t = \boldsymbol{\omega}_{i/e}^t = \omega_{i/e}[\cos\bar{\phi}, 0, -\sin\bar{\phi}]^\intercal$ where $\bar{\phi}$ is the latitude and

$$\omega_{i/e} \approx 7.292115 \times 10^{-5} \frac{rad}{sec}$$

is the Earth rotation rate.

To represent the attitude of the rover in tangent frame, the Euler angle with a rotation sequence z-y-x is used [22]. Let $\boldsymbol{q}_t^b = [\phi, \theta, \psi]^\intercal$ be the roll-pitch-yaw vector, then the rotation matrix $\boldsymbol{R}_t^b \in SO(3)$ from the tangent frame to the body frame is

$$\boldsymbol{R}_t^b = \boldsymbol{R}_1(\phi)\boldsymbol{R}_2(\theta)\boldsymbol{R}_3(\psi), \tag{3.10}$$

---

[4]$\dot{\boldsymbol{\Omega}}_{i/t}^t \equiv \boldsymbol{0}$, since the Earth rotation rate is assumed constant.

where

$$\boldsymbol{R}_1(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c\phi & s\phi \\ 0 & -s\phi & c\phi \end{bmatrix}, \boldsymbol{R}_2(\theta) = \begin{bmatrix} c\theta & 0 & -s\theta \\ 0 & 1 & 0 \\ s\theta & 0 & c\theta \end{bmatrix}, \boldsymbol{R}_3(\psi) = \begin{bmatrix} c\psi & s\psi & 0 \\ -s\psi & c\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

To obtain the derivative of the Euler angle $\dot{\boldsymbol{q}}_t^b = [\dot{\phi}, \dot{\theta}, \dot{\psi}]^\intercal$, it should be noted that $\dot{\boldsymbol{q}}_t^b$ is not related to the angular rate $\boldsymbol{\omega}_{t/b}^b$ through only one rotation. This is because each of (ϕ, θ, ψ) and then $(\dot{\phi}, \dot{\theta}, \dot{\psi})$ is defined in a different reference frame [22]. Instead, $\boldsymbol{\omega}_{t/b}^b$ can be divided into 3 components corresponding to 3 rotations in the z-y-x sequence,

$$\boldsymbol{\omega}_{t/b}^b = \boldsymbol{\omega}_{t/1}^b + \boldsymbol{\omega}_{1/2}^b + \boldsymbol{\omega}_{2/b}^b, \tag{3.11}$$

where

$$\boldsymbol{\omega}_{t/1}^b = \boldsymbol{R}_1(\phi)\boldsymbol{R}_2(\theta) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \dot{\psi}, \quad \boldsymbol{\omega}_{1/2}^b = \boldsymbol{R}_1(\phi) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \dot{\theta}, \quad \boldsymbol{\omega}_{2/b}^b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \dot{\phi}.$$

Thus, eqn. (3.11) yields

$$\boldsymbol{\omega}_{t/b}^b = \begin{bmatrix} 1 & 0 & -s\theta \\ 0 & c\phi & s\phi c\theta \\ 0 & -s\phi & c\phi c\theta \end{bmatrix} \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix}, \tag{3.12}$$

41

and then

$$\dot{\boldsymbol{q}}_t^b = \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -s\theta \\ 0 & c\phi & s\phi c\theta \\ 0 & -s\phi & c\phi c\theta \end{bmatrix}^{-1} \boldsymbol{\omega}_{t/b}^b \triangleq \boldsymbol{\Omega}_E^{-1} \boldsymbol{\omega}_{t/b}^b. \tag{3.13}$$

In eqn. (3.13), when $\theta = \pm\pi/2$, i.e., pitch is $\pm 90°$,

$$\boldsymbol{\Omega}_E^{-1} = \begin{bmatrix} 1 & s\phi \cdot t\theta & c\phi \cdot t\theta \\ 0 & c\phi & -s\phi \\ 0 & s\phi/c\theta & c\phi/c\theta \end{bmatrix}$$

reaches singularity. For the navigation scenarios (i.e., underwater, land vehicle) focused on in this dissertation, the pitch angle is expect to rarely reach $\pm 90°$. So, in this dissertation the Euler angle representation is proper for use. For other cases where the pitch angle can reach $\pm 90°$, alternative representations, e.g. quaternion [74], can be used to approach the singularity.

With eqn. (3.9), (3.13) and $\boldsymbol{\omega}_{t/b} = \boldsymbol{\omega}_{i/b} - \boldsymbol{\omega}_{i/t}$, the continuous-time kinematic model is

$$\left. \begin{aligned} \dot{\boldsymbol{p}}_{t/b}^t &= \boldsymbol{v}_{t/b}^t \\ \dot{\boldsymbol{v}}_{t/b}^t &= \boldsymbol{R}_b^t \mathbf{f}_{i/b}^b + \mathbf{g}_{i/b}^t - 2\boldsymbol{\Omega}_{i/t}^t \boldsymbol{v}_{t/b}^t \\ \dot{\boldsymbol{q}}_t^b &= \boldsymbol{\Omega}_E^{-1} \left(\boldsymbol{\omega}_{i/b}^b - \boldsymbol{R}_t^b \boldsymbol{\omega}_{i/e}^t \right) \end{aligned} \right\}. \tag{3.14}$$

3.1.3 Inertial Measurements

As a high-rate sensor, Inertial Measurement Unit (IMU) give samples of the specific force vector (see eqn. (3.8)) via an accelerometer and the angular rate of rotation

via a gyroscope,

$$\tilde{\boldsymbol{y}}_a = \mathbf{f}_{i/b}^b + \boldsymbol{b}_a + \boldsymbol{n}_a, \tag{3.15}$$

$$\tilde{\boldsymbol{y}}_g = \boldsymbol{\omega}_{i/b}^b + \boldsymbol{b}_g + \boldsymbol{n}_g, \tag{3.16}$$

where $\boldsymbol{b}_a$, $\boldsymbol{b}_g$ are bias vectors and $\boldsymbol{n}_a$, $\boldsymbol{n}_g$ are noise vectors. It is conventional to assume that $\boldsymbol{n}_a$ and $\boldsymbol{n}_g$ are white Gaussian processes and

$$\boldsymbol{n}_a \sim \mathcal{N}(\mathbf{0}, \sigma_a^2 \boldsymbol{I}), \quad \boldsymbol{n}_g \sim \mathcal{N}(\mathbf{0}, \sigma_g^2 \boldsymbol{I}).$$

Furthermore, the bias $\boldsymbol{b}_a$ and $\boldsymbol{b}_g$ are often modeled with random walks as

$$\dot{\boldsymbol{b}}_a = \boldsymbol{v}_a, \quad \ddot{\boldsymbol{b}}_a = \mathbf{0}, \tag{3.17}$$

$$\dot{\boldsymbol{b}}_g = \boldsymbol{v}_g, \quad \ddot{\boldsymbol{b}}_g = \mathbf{0}, \tag{3.18}$$

where the driving noise $\boldsymbol{v}_a$ and $\boldsymbol{v}_g$ are also assumed white Gaussian,

$$\boldsymbol{v}_a \sim \mathcal{N}(\mathbf{0}, \sigma_{ba}^2 \boldsymbol{I}), \quad \boldsymbol{v}_g \sim \mathcal{N}(\mathbf{0}, \sigma_{bg}^2 \boldsymbol{I}).$$

Thus, with the IMU measurements and the estimated bias vectors, the specific force and the angular rate can be evaluated as

$$\hat{\mathbf{f}}_{i/b}^b = \tilde{\boldsymbol{y}}_a - \hat{\boldsymbol{b}}_a, \quad \hat{\boldsymbol{\omega}}_{i/b}^b = \tilde{\boldsymbol{y}}_g - \hat{\boldsymbol{b}}_g. \tag{3.19}$$

3.1.4   INS Propagation

By defining the INS state vector

$$\boldsymbol{x} = \left[\left(\boldsymbol{p}_{t/b}^t\right)^{\mathsf{T}} \quad \left(\boldsymbol{v}_{t/b}^t\right)^{\mathsf{T}} \quad \left(\boldsymbol{q}_t^b\right)^{\mathsf{T}} \quad \boldsymbol{b}_a^{\mathsf{T}} \quad \boldsymbol{b}_g^{\mathsf{T}} \right]^{\mathsf{T}}, \tag{3.20}$$

43

as well as the input vector

$$\boldsymbol{u} = \left[\left(\mathbf{f}_{i/b}^b \right)^\mathsf{T} \quad \left(\boldsymbol{\omega}_{i/b}^b \right)^\mathsf{T} \right]^\mathsf{T}, \tag{3.21}$$

the continuous-time model in eqn. (3.1) can be specified as

$$
\left.
\begin{aligned}
\dot{\boldsymbol{p}}_{t/b}^t &= \boldsymbol{v}_{t/b}^t \\[4pt]
\dot{\boldsymbol{v}}_{t/b}^t &= \boldsymbol{R}_b^t \left(\tilde{\boldsymbol{y}}_a - \boldsymbol{b}_a - \boldsymbol{n}_a \right) + \mathbf{g}_{i/b}^t - 2\boldsymbol{\Omega}_{i/t}^t \boldsymbol{v}_{t/b}^t \\[4pt]
\dot{\boldsymbol{q}}_t^b &= \boldsymbol{\Omega}_E^{-1} \left(\tilde{\boldsymbol{y}}_g - \boldsymbol{b}_g - \boldsymbol{n}_g - \boldsymbol{R}_t^b \boldsymbol{\omega}_{i/e}^t \right) \\[4pt]
\dot{\boldsymbol{b}}_a &= \boldsymbol{v}_a \\[4pt]
\dot{\boldsymbol{b}}_g &= \boldsymbol{v}_g
\end{aligned}
\right\} . \tag{3.22}
$$

In operations, INS keeps propagating the state estimate

$$\hat{\boldsymbol{x}} = \left[\left(\hat{\boldsymbol{p}}_{t/b}^t \right)^\mathsf{T} \quad \left(\hat{\boldsymbol{v}}_{t/b}^t \right)^\mathsf{T} \quad \left(\hat{\boldsymbol{q}}_t^b \right)^\mathsf{T} \quad \hat{\boldsymbol{b}}_a^\mathsf{T} \quad \hat{\boldsymbol{b}}_g^\mathsf{T} \right]^\mathsf{T}, \tag{3.23}$$

through numerical integration of

$$
\left.
\begin{aligned}
\dot{\hat{\boldsymbol{p}}}_{t/b}^t &= \dot{\hat{\boldsymbol{v}}}_{t/b}^t \\[4pt]
\dot{\hat{\boldsymbol{v}}}_{t/b}^t &= \hat{\boldsymbol{R}}_b^t \left(\tilde{\boldsymbol{y}}_a - \hat{\boldsymbol{b}}_a \right) + \hat{\mathbf{g}}_{i/b}^t - 2\hat{\boldsymbol{\Omega}}_{i/t}^t \hat{\boldsymbol{v}}_{t/b}^t \\[4pt]
\dot{\hat{\boldsymbol{q}}}_t^b &= \hat{\boldsymbol{\Omega}}_E^{-1} \left(\tilde{\boldsymbol{y}}_g - \hat{\boldsymbol{b}}_g - \hat{\boldsymbol{R}}_t^b \hat{\boldsymbol{\omega}}_{i/e}^t \right) \\[4pt]
\dot{\hat{\boldsymbol{b}}}_a &= \boldsymbol{0} \\[4pt]
\dot{\hat{\boldsymbol{b}}}_g &= \boldsymbol{0}
\end{aligned}
\right\}, \tag{3.24}
$$

where the local gravity vector $\hat{\mathbf{g}}_{i/b}^t(\hat{\boldsymbol{p}}_{t/b}^t)$ and the tangent frame rotation rate $\hat{\boldsymbol{\Omega}}_{i/t}^t(\hat{\boldsymbol{p}}_{t/b}^t)$ are functions of the estimated position. By rewriting eqn. (3.22) and (3.24) compactly, the eqn. (3.1) and (3.2) can be derived.

Let $\delta\boldsymbol{x}$ denote the INS *error state* from the estimated state to the true system state,

$$\delta\boldsymbol{x} = \left[\left(\delta\boldsymbol{p}_{t/b}^t\right)^\mathsf{T} \quad \left(\delta\boldsymbol{v}_{t/b}^t\right)^\mathsf{T} \quad (\delta\boldsymbol{\theta})^\mathsf{T} \quad \delta\boldsymbol{b}_a^\mathsf{T} \quad \delta\boldsymbol{b}_g^\mathsf{T}\right]^\mathsf{T} \in \mathbb{R}^{n_s}, \tag{3.25}$$

where $n_s \in \mathbb{Z}^+$ is the dimension of the INS error state, and $\delta\boldsymbol{\theta}$ represent the small-angle rotation from the estimate body frame to true body frame such that

$$\boldsymbol{R}_b^t = \left(\boldsymbol{I}_3 - [\delta\boldsymbol{\theta}\times]\right)\hat{\boldsymbol{R}}_b^t. \tag{3.26}$$

For any initial state $\boldsymbol{x}(\tau_k)$, the solution to (3.1) for $t \in [\tau_k, \tau_{k+1}]$ is

$$\boldsymbol{x}(t) = \boldsymbol{x}(\tau_k) + \int_{\tau_k}^t \boldsymbol{f}(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau))d\tau. \tag{3.27}$$

While nature solves (3.27) in continuous time, the INS only has IMU and aiding measurements at discrete time instants; therefore, the INS numerically solves

$$\hat{\boldsymbol{x}}(\tau_{k+1}) = \boldsymbol{\phi}\big(\hat{\boldsymbol{x}}(\tau_k), \tilde{\boldsymbol{u}}(\tau_k)\big)$$

$$= \hat{\boldsymbol{x}}(\tau_k) + \int_{\tau_k}^{\tau_{k+1}} \boldsymbol{f}(\hat{\boldsymbol{x}}(\tau), \tilde{\boldsymbol{u}}(\tau))d\tau, \tag{3.28}$$

where $\boldsymbol{\phi}$ is defined as the integration operator. The result of the numeric integration of (3.28) is the INS state estimate of $\hat{\boldsymbol{x}}(\tau_{k+1})$ given $\hat{\boldsymbol{x}}(\tau_k)$ and $\hat{\boldsymbol{u}}(\tau_k)$. The numeric integration repeats to propagate the state measurements between the times of aiding measurements. The aiding measurement times can be unequally spaced in time without causing any complications.

Let $\tilde{\mathbf{U}}_j = \{\tilde{\boldsymbol{u}}(\tau_k), \ \tau_k \in [t_j, t_{j+1}]\}$, then eqn. (3.28) can be called recursively to compute $\hat{\boldsymbol{x}}(t_{j+1})$ from $\hat{\boldsymbol{x}}(t_j)$ and $\tilde{\mathbf{U}}_j$, denote this as

$$\hat{\boldsymbol{x}}(t_{j+1}) = \boldsymbol{\phi}\big(\hat{\boldsymbol{x}}(t_j), \tilde{\mathbf{U}}_j\big). \tag{3.29}$$

At the same time, nature is integrating eqn. (3.27) which it denoted as

$$\boldsymbol{x}(t_{j+1}) = \boldsymbol{\phi}\big(\boldsymbol{x}(t_j), \mathbf{U}_j\big). \tag{3.30}$$

The linearized error growth model is

$$\delta\hat{\boldsymbol{x}}(t_{j+1}) = \boldsymbol{\Phi}_j \delta\hat{\boldsymbol{x}}(t_j) + \boldsymbol{w}_j \tag{3.31}$$

where $\boldsymbol{\omega}_j \sim N(0, \mathbf{Q}_j)$ is the discretized noise due to the existence of $[\boldsymbol{n}_a \quad \boldsymbol{n}_g \quad \boldsymbol{v}_a \quad \boldsymbol{v}_g]$, and $\boldsymbol{\Phi}_j$ is the error state transition matrix.

The INS provides both $\mathbf{Q}_j$ and $\boldsymbol{\Phi}_j$, for example see Section 7.2.5.2 in [22]. In [44], other popular methods computing the error transition matrix $\boldsymbol{\Phi}_j$ are well discussed, and furthermore a closed-form solution is presented.

3.1.5   Extended Kalman Filtering based Aided INS

When (standard) aiding measurements of the form

$$\tilde{\boldsymbol{z}}(t) = \boldsymbol{h}(\boldsymbol{x}(t)) + \boldsymbol{n}_z(t), \quad \boldsymbol{n}_z \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_z), \tag{3.32}$$

are available, various methods are available to use the initial state, inertial measurements, and aiding measurement information to estimate the vehicle state vector

[7, 22, 37]. Extended Kalman Filter (EKF) is widely applied as the estimator of the aided INS systems, due to its simplicity for implementation and the real-time efficiency.

The standard EKF can be revisited in a Weighted Least Square (WLS) form [22, 27]. Given the INS prior and the aiding measurements at time step t_k,

$$\hat{\boldsymbol{x}}_k^- = \boldsymbol{x}_k + \delta\boldsymbol{x}_k^-, \quad \delta\boldsymbol{x}_k^- \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_\mathbf{k}^-), \tag{3.33}$$

$$\tilde{\boldsymbol{z}}_k = \boldsymbol{h}_k(\boldsymbol{x}_k) + \boldsymbol{n}_k, \quad \boldsymbol{n}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_\mathbf{k}), \tag{3.34}$$

as well as the standard White-Gaussian-Noise (WGN) assumption on the process noise $\boldsymbol{w}_k$ and the measurement noise $\boldsymbol{n}_k$ [7, 22, 37], a *Maximum Likelihood* estimation of the state correction $\delta\boldsymbol{x}_k$ at t_k can be formulated as,

$$\delta\boldsymbol{x}_k^+ = \underset{\delta\boldsymbol{x}_k}{arg\,max} \ p_{\delta\boldsymbol{x}_k^-}(\delta\boldsymbol{x}_k)p_{\boldsymbol{n}_k}(\tilde{\boldsymbol{z}}_k - \mathbf{H}_k\hat{\boldsymbol{x}}_k^- - \mathbf{H}_k\delta\boldsymbol{x}_k), \tag{3.35}$$

where $\mathbf{H}_k = \frac{\partial\boldsymbol{h}_k}{\partial\boldsymbol{x}}\big|_{\boldsymbol{x}=\hat{\boldsymbol{x}}_k^-}$ is the Jacobian matrix of the measurement model evaluated at the INS prior $\hat{\boldsymbol{x}}_k^-$.

Furthermore, the following Least Square problem can be derived by evaluating the negative log-likelihood of the right hand side of eqn. (3.36),

$$\delta\boldsymbol{x}_k^+ = \underset{\delta\boldsymbol{x}_k}{arg\,min} \ \|\delta\boldsymbol{x}_k\|_{\mathbf{P}_\mathbf{k}^-}^2 + \|\delta\boldsymbol{z}_k - \mathbf{H}_k\delta\boldsymbol{x}_k\|_{\mathbf{R}_\mathbf{k}}^2, \tag{3.36}$$

where $\delta\boldsymbol{z}_k \triangleq \tilde{\boldsymbol{z}}_k - \mathbf{H}_k\hat{\boldsymbol{x}}_k^-$ is the measurement residual (a.k.a. innovation) and the notation $\|\boldsymbol{x}\|_\mathbf{C}^2 \triangleq \boldsymbol{x}^\intercal\mathbf{C}^{-1}\boldsymbol{x}$ is defined as the Mahalanobis distance.

Letting

$$
\mathbf{A}_k = \begin{bmatrix} \mathbf{I} \\ \mathbf{H}_k \end{bmatrix}, \quad \mathbf{C}_k = \begin{bmatrix} \mathbf{P}_\mathbf{k}^- & \mathbf{0} \\ \mathbf{0} & \mathbf{R_k} \end{bmatrix} \quad \text{and} \quad \delta \mathbf{Z}_k = \begin{bmatrix} \mathbf{0} \\ \delta \mathbf{z}_k \end{bmatrix},
$$

then the Kalman Filter measurement update can be derived by solving the Weighted Least Square,

$$
\min_{\delta \boldsymbol{x}} \; \| \delta \mathbf{Z}_k - \mathbf{A}_k \delta \boldsymbol{x} \|_{\mathbf{C}_k}. \tag{3.37}
$$

The solution of eqn. (3.37) is

$$
\delta \boldsymbol{x}_k^+ = \left(\mathbf{A}_k^\intercal \mathbf{C}_k^{-1} \mathbf{A}_k \right)^{-1} \mathbf{A}_k^\intercal \mathbf{C}_k^{-1} \delta \mathbf{Z}_k, \tag{3.38}
$$

and the corresponding covariance of $\delta \boldsymbol{x}_k^+$ is $\mathbf{P}_k^+ = \left(\mathbf{A}_k^\intercal \mathbf{C}_k^{-1} \mathbf{A}_k \right)^{-1}$. With $\delta \boldsymbol{x}_k^+$, the INS state is updated as

$$
\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k^- + \delta \boldsymbol{x}_k^+,
$$

and the renowned Kalman gain

$$
\boldsymbol{K}_k = \left(\mathbf{A}_k^\intercal \mathbf{C}_k^{-1} \mathbf{A}_k \right)^{-1} \mathbf{A}_k^\intercal \mathbf{C}_k^{-1} = \mathbf{P}_\mathbf{k}^- \mathbf{H}_\mathbf{k}^\intercal \left(\mathbf{R_k} + \mathbf{H}_k \mathbf{P}_\mathbf{k}^- \mathbf{H}_k^\intercal \right)
$$

can also be derived by expanding $\left(\mathbf{A}_k^\intercal \mathbf{C}_k^{-1} \mathbf{A}_k \right)^{-1} \mathbf{A}_k^\intercal \mathbf{C}_k^{-1}$.

EKF works as a standard framework for many inertial navigation application, e.g. GPS-INS [22], Underwater-INS [53], Vision-Inertial-Odemetry (VIO) [43–49, 55]. However, the performance of the EKF significantly depends on initial conditions and nonlinearities (see [18]). This is due to the fact that previous improper EKF
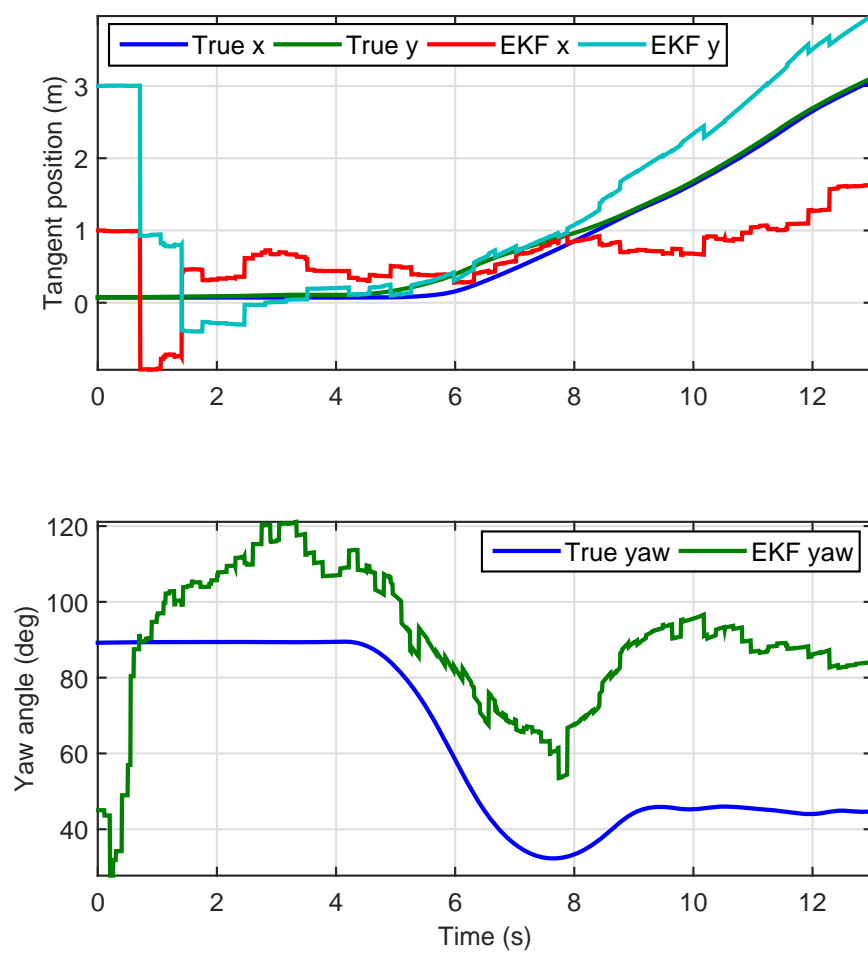
Figure 3.1: Divergence of EKF under poor initialization.

linearization points cannot be corrected at later times. Fig. (3.1) shows the divergence of EKF when the initial yaw is set wrong.

To overcome this challenge of the EKF method, a smoothing approach has attracted considerable attention in the Simultaneous Localization and Mapping (SLAM) research community (see [18, 34, 35]). In Section 3.2, the graph optimization based smoothing method is extended to INS, in order to enhance the navigation performance in accuracy and reliability.

3.2 Graph Modeling for Inertial Navigation

A key point of smoothing in the SLAM context is keeping the complete robot trajectory in the estimation, so all the useful information over the time window can be considered to reach statistically optimality. Nonlinear optimization is considered and solved iteratively. Thus, the previous linearization point of the system and the measurement models can be corrected in later runs, when new data comes. Although the computational cost is higher than EKF, by exploring the sparsity of the involved matrices and introducing incremental solution, the smoothing approach is claimed to be fast and efficient for practical applications (see [35]).

In this work, the smoothing method is extended to Inertial Navigation in order to enhance the navigation performance in accuracy and reliability. Without loss of generality, the problem stated as follows is considered in each smoothing window.

Problem 3.1. *For a system described by eqn. (3.1), to estimate the trajectory of the rover* $\mathbf{X} = [\boldsymbol{x}^{\mathsf{T}}(t_0), \ldots, \boldsymbol{x}^{\mathsf{T}}(t_K)]^{\mathsf{T}} \in \mathbb{R}^{n_s(K+1)}$ *over the time interval* $[t_0, t_K]$*, we have*

- *an initial distribution for the state* $\boldsymbol{x}(t_0) \sim \mathcal{N}(\boldsymbol{x}_0, \mathbf{P}_0)$,

- *IMU measurements* $\mathbf{U} = \{\mathbf{U}_k\}_{k=0}^{K-1}$*, where*

$$\mathbf{U}_k = \{\tilde{\boldsymbol{u}}(\tau_n), t_k \leq \tau_n \leq t_{k+1}\},$$

- *the aiding measurements* $\mathbf{Z} = \{\tilde{\boldsymbol{z}}_k\}_{k=1}^{K}$*, where* $\boldsymbol{z}_k$ *has a general model as*

$$\tilde{\boldsymbol{z}}_k(t_k) = \boldsymbol{h}(\mathbf{X}(t_k)) + \boldsymbol{n}_z(t_k), \quad \boldsymbol{n}_z \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R}_{z_k}), \tag{3.39}$$

where t_0*,* $t_k \in (\tau_0, \tau_\kappa]$ *and the set* $\{\tau_n\}$ *contains the high frequency IMU measurement time instants.* $\triangle$

Note that at time step t_k, $\tilde{\boldsymbol{z}}_k(t_k)$ can be a measurement depending on multiple previous states $\{\boldsymbol{x}(t_0), \ldots, \boldsymbol{x}(t_k)\}$. This measurement model assumption is different from that of standard EKF which only handle the measurements of current state.

Then, the objective is:

Objective 3.2. *Estimate the optimal state trajectory* $\mathbf{X} = [\boldsymbol{x}^{\mathsf{T}}(t_0), \ldots, \boldsymbol{x}^{\mathsf{T}}(t_K)]^{\mathsf{T}} \in \mathbb{R}^{n_s(K+1)}$*with the given sensor measurements* $\mathbf{U}$*,* $\mathbf{Z}$ *and the prior state density* $p_{\boldsymbol{x}}(\boldsymbol{x}(t_0))$*.* $\triangle$

Under Kalman Filtering frame, only the current state and uncertainty are maintained at each time step. In smoothing, when a time window of navigation data is

considered, the formulation of the estimation problem may be complicated to handle in practice. The state-of-art approach is to divide the smoothing framework into a front-end representing the estimation formulation by a probabilistic graphical model and a back-end converting the graph into matrix form and executing optimization.

In the past decade, factor graph modeling has attracted considerable attention in signal processing, artificial intelligence, and robotics as a means to represent complex Bayesian estimation problems [40, 50], e.g. Simultaneous Localization and Mapping (SLAM) [17]. Often times, *Maximum-a-Posteriori* (MAP) estimates can be derived from factor graphs. In MAP estimation, all the information (e.g., prior, kinematics, sensor data) over a time window is considered, and then iterative nonlinear optimization is applied to achieve statistical optimality. Several efficient software toolboxes of graph optimization based estimation have been developed in robotics and computer science community, e.g. g2o [41], GTSAM [17], MTK [79].

To illustrate the factor graph modeling idea, first a Bayes Network (BN) [39] is shown in Fig. 3.2 to represent the inertial navigation problem defined above. For simplicity, only states $\{\boldsymbol{x}(t_0), \boldsymbol{x}(t_3)\}$ corresponding measurements are shown explicitly. Note that $\boldsymbol{z}(t_1)$ and $\boldsymbol{z}(t_2)$ are measurements of just the state at one time step, but $\boldsymbol{z}(t_3)$ is measurement of two states. There can be measurements depending more three or more states in the trajectory. The Bayes Network is a directed acyclic graph representing the joint distribution $p(\mathbf{X}, \mathbf{Z}, \mathbf{U})$. Each node in the BN represents a variable and its associated conditional density. For example, in Fig. 3.2 the node

labeled $\boldsymbol{x}(t_1)$ represents the conditional probability of the unknown state $\boldsymbol{x}(t_1)$ given $\boldsymbol{x}(t_0)$ and $\mathbf{U}_0$. So, $p(\mathbf{X}, \mathbf{Z}, \mathbf{U})$ can be regarded as 'factorized' over the BN by following Bayes rule. This representation motivates another type of graphical model, named *Factor Graph* shown in Fig. 5.3. In a factor graph, only the unknown variables are explicitly labeled as nodes, while the edges (with black dots) represents the 'factors' which encode the probabilistic information (constraints) between the unknown variables. In particular, in Fig. 5.3 the factor above the node $\boldsymbol{x}(t_0)$ is for the initial state prior. The factors between two consecutive purple nodes are the probabilistic constraints derived from the IMU data set and kinematic constraints. The factors above the nodes $\{\boldsymbol{x}(t_k)\}_{k=1}^3$ are for the aiding measurements. The factors can be unary, e.g. prior knowledge and $\boldsymbol{z}(t_1)$, $\boldsymbol{z}(t_2)$ measurements, or be binary, e.g. INS constraints and $\boldsymbol{z}(t_3)$ measurements.
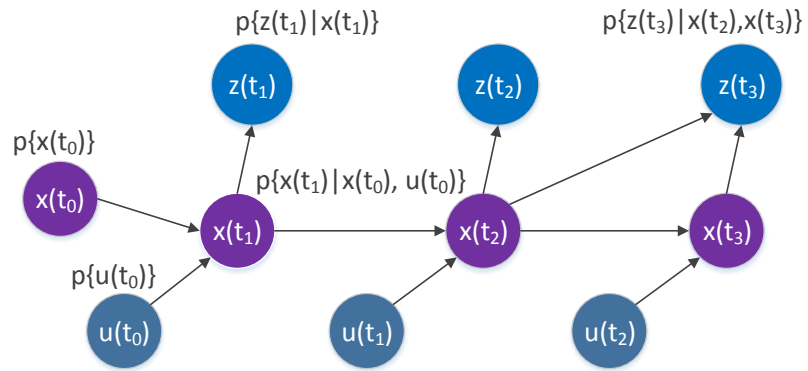
Figure 3.2: Bayes Network representing general INS smoothing problem
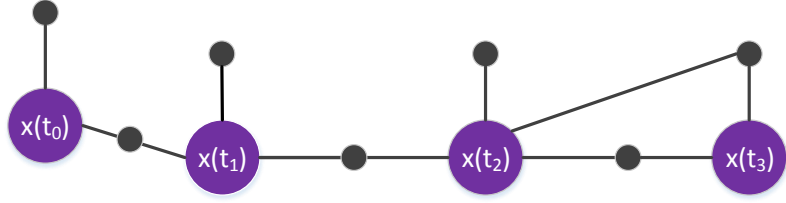
Figure 3.3: The INS navigation estimation problem over a time window represented by a Factor Graph. The labeled nodes are for the unknown variables, and the black dotted edges are the factors encoding the probabilistic information between the unknown variables.

Factor graphs have been widely used to represent and solve complex estimation problems, e.g. SLAM. Existing factor graph based C++ libraries [17, 41, 79] can facilitate the formulation of complex estimation problem significantly, by maintaining a factor graph in the front-end. Thus, including a new unknown variable or a new measurement into the estimation problem is accomplished by simply adding a node or a edge into the proper place of the current graph. These libraries also provide robust and efficient solutions as a 'back-end' for the estimation problems represented by the factor graph. In the factor graph based estimation solution, nonlinear least square methods such as Levenberg-Marquardt or Powell's Dog-Leg are applied and the sparsity nature of the graph is exploited to achieve computational efficiency.

With the aid of the factor graph in Fig. 5.3, the following relation can be derived

$$p(\mathbf{X}, \mathbf{U}, \mathbf{Z})$$

$$= p(\mathbf{X}_+, \boldsymbol{x}(t_0), \mathbf{U}, \mathbf{Z}) \tag{3.40}$$

$$= p(\mathbf{X}_+, \boldsymbol{x}(t_0), \mathbf{U})p(\mathbf{Z}|\mathbf{X}_+, \boldsymbol{x}(t_0), \mathbf{U}) \tag{3.41}$$

$$= p(\mathbf{X}_+, \boldsymbol{x}(t_0), \mathbf{U})p(\mathbf{Z}|\mathbf{X}) \tag{3.42}$$

$$= p(\boldsymbol{x}(t_0), \mathbf{U})p(\mathbf{X}_+|\boldsymbol{x}(t_0), \mathbf{U})p(\mathbf{Z}|\mathbf{X}) \tag{3.43}$$

$$= p(\boldsymbol{x}(t_0))p(\mathbf{X}_+|\boldsymbol{x}(t_0), \mathbf{U})p(\mathbf{Z}|\mathbf{X}), \tag{3.44}$$

where $\mathbf{X}_+ = \{\boldsymbol{x}(t) \text{ for } t = t_1, \ldots, t_K\}$. The eqn. (3.40) is from separating $\mathbf{X}$ into $\boldsymbol{x}(t_0)$ and $\mathbf{X}_+$. The eqn. (3.41) and (3.43) are from the definition of conditional probability. The eqn. (3.42) is from the standard assumption that $\mathbf{U}$, $\mathbf{Z}$ (i.e., measurements from different sensors) are measured independently. The eqn. (3.44) derives from that the fact that the prior $\boldsymbol{x}(t_0)$ is independent of the IMU measurements $\mathbf{U}$.

Thus, a *Maximum-a-Posteriori* can be formed to estimate $\mathbf{X}$ by maximizing the right hand side of eqn. (6.2),

$$\max_{\mathbf{X} \in \mathbb{R}^{n_s(K+1)}, \mathbf{N} \in \mathbb{Z}^m} p(\boldsymbol{x}(t_0))p(\mathbf{X}_+|\boldsymbol{x}(t_0), \mathbf{U})p(\mathbf{Z}|\mathbf{X}). \tag{3.45}$$

With the Gaussian noise assumption, the negative log-likelihood of the right hand side of eqn. (3.44) can be rewritten as

$$
\begin{aligned}
\|\mathbf{v}\|_{\mathbf{W}}^2 = &\|\boldsymbol{x}(t_0) - \boldsymbol{x}_0\|_{\mathbf{P}_0}^2 \\
&+ \sum_k \|\boldsymbol{\phi}\big(\boldsymbol{x}(t_k), \mathbf{U}_k\big) - \boldsymbol{x}(t_{k+1})\|_{\mathbf{Q}_k}^2 \\
&+ \sum_k \|\boldsymbol{h}(\mathbf{X}(t_k)) - \boldsymbol{z}(t_k)\|_{\boldsymbol{R}_{z_k}}^2
\end{aligned}
\tag{3.46}
$$

where $\|\mathbf{v}\|_{\mathbf{W}}^2 = \mathbf{v}^\top \mathbf{W}^{-1} \mathbf{v}$ is the squared Mahalanobis distance with the matrix $\mathbf{W}$. All terms on the right-hand side also use this notation. The vector $\mathbf{v}$ is the concatenation of each of the vectors summed in the right-hand side of eqn. (6.4). In eqn. (6.4), the operator $\boldsymbol{\phi}$ and the covariance matrix $\mathbf{Q}_k$ are defined in the INS review portion of the appendix. The matrix $\mathbf{W}$ is the positive definite block diagonal matrix formed by the positive definite submatrices $\mathbf{Q}_k$, $\mathbf{P}_0$ and $\boldsymbol{R}_{z_k}$. The IMU data set $\mathbf{U}_k$ contains all IMU measurements in the interval $[t_k, t_{k+1}]$.

Let $\boldsymbol{\Sigma}_W^\top \boldsymbol{\Sigma}_W = \mathbf{W}^{-1}$, then $\boldsymbol{r} \triangleq \boldsymbol{\Sigma}_W \mathbf{v}$ is the weighted residual and $\|\mathbf{v}\|_{\mathbf{W}}^2 = \|\boldsymbol{r}\|^2$. So, the MAP estimation is transformed into a Nonlinear Least Square (NLS) problem,

$$
\mathbf{X}^* = \underset{\mathbf{X} \in \mathbb{R}^{ns(K+1)}}{arg\,min} \|\boldsymbol{r}(\mathbf{X})\|^2,
\tag{3.47}
$$

where $\boldsymbol{r}$ is a vector representing the factors in the graph model:

$$\boldsymbol{r}(\mathbf{X},\mathbf{N}) = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{P}_0}\left(\boldsymbol{x}(t_0) - \boldsymbol{x}_0\right) \\[6pt] \boldsymbol{\Sigma}_{\mathbf{Q}_0}\left(\boldsymbol{\phi}\big(\boldsymbol{x}(t_0), \mathbf{U}_0\big) - \boldsymbol{x}(t_1)\right) \\[6pt] \vdots \\[6pt] \boldsymbol{\Sigma}_{\mathbf{Q}_{K-1}}\left(\boldsymbol{\phi}\big(\boldsymbol{x}(t_{K-1}), \mathbf{U}_{K-1}\big) - \boldsymbol{x}(t_K)\right) \\[6pt] \boldsymbol{\Sigma}_{\boldsymbol{R}_{z_1}}\left(\boldsymbol{h}(\mathbf{X}(t_1)) - \boldsymbol{z}(t_1)\right) \\[6pt] \vdots \\[6pt] \boldsymbol{\Sigma}_{\boldsymbol{R}_{z_K}}\left(\boldsymbol{h}(\mathbf{X}(t_K)) - \boldsymbol{z}(t_K)\right) \end{bmatrix}. \tag{3.48}$$

In the following, a Gauss-Newton method is presented as an example for solving the NLS derived from the graph model.

3.3 Graph Optimization for Inertial Navigation

This section presents the solution to solve the graph optimization which formed in Section 3.2 for inertial navigation.

3.3.1 Notation

Herein useful notation is specified for later analysis. Some of the notation is similar to that used in [41]. Let $\mathcal{M} \subset \mathbb{R}^2$ contain all tuples $(i, j) \subset \{0, \ldots, K\} \times \{0, \ldots, K\}$ such that there exist measurement(s) depending on $\boldsymbol{x}(t_i)$ and $\boldsymbol{x}(t_j)$. Herein, for simplicity of illustration, it is assumed that each measurement only depends on up

to TWO states in the trajectory. The presented method can be extended to more complicated cases straightforwardly. Note that $(i, i) \in \mathcal{M}$ is also valid for standard measurement(s) received at t_i or $i = 0$ (for the prior). Let $m_{ij} \in \mathbb{Z}^+$ denote the total number of (scalar) measurements associated to $(i, j) \in \mathcal{M}$. With $\boldsymbol{x}_i \triangleq \boldsymbol{x}(t_i)$ and $\boldsymbol{x}_j \triangleq \boldsymbol{x}(t_j)$, the vector $\boldsymbol{\nu}_{ij}(\boldsymbol{x}_i, \boldsymbol{x}_j) \in \mathbb{R}^{m_{ij}}$ concatenates the residuals of all measurements depending on $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, and $\mathbf{R}_{ij} \in \mathbb{R}^{m_{ij} \times m_{ij}}$ is the corresponding covariance matrix for $\boldsymbol{\nu}_{ij}(\boldsymbol{x}_i, \boldsymbol{x}_j)$ computed from $\mathbf{P}_0$, $\mathbf{Q}_k$ and $\mathbf{R}_{z_k}$.

3.3.2 Solution Overview

The algorithm will involve two steps.

1. Given an estimate of the solution

$$\hat{\mathbf{X}}^l = \left\{ \hat{\boldsymbol{x}}^l(t) \text{ for } t = t_0, \ldots, t_K \right\},$$

which is treated as a vector in $\mathbb{R}^{n(K+1)}$, an optimization algorithm computes a $\delta \mathbf{X}^l \in \mathbb{R}^{n(K+1)}$ such that

$$\hat{\mathbf{X}}^{l+1} = \hat{\mathbf{X}}^l + \delta \mathbf{X}^l \tag{3.49}$$

is an improved solution to eqn. (6.3).

2. Evaluation of the factor $p_{\boldsymbol{\omega}_u}\left(\mathbf{X}_+ - \boldsymbol{\phi}(\mathbf{X}, \mathbf{U})\right)$ of the MAP cost function and its linearization requires reintegration relative to $\hat{\mathbf{X}}^{l+1}$. This issue is subtle and will be clarified in Subsection 3.3.3.

3.3.3   Solution Detail: Reintegration

In this subsection, to simplify notation, the subscript k is used to represent the time t_k, with $t_k = t_0, \ldots, t_K$. For example, $\hat{\boldsymbol{x}}_k^l = \hat{\boldsymbol{x}}^l(t_k)$.

Given $\hat{\mathbf{X}}^l$, the kinematic model provides an operator (see eqn. (3.28))

$$\bar{\boldsymbol{x}}_{k+1}^l = \hat{\boldsymbol{x}}_k^l + \int_{t_k}^{t_{k+1}} f(\bar{\boldsymbol{x}}^l(\tau), \hat{\boldsymbol{u}}(\tau)) d\tau \tag{3.50}$$

with $\bar{\boldsymbol{x}}^l(t_k) = \hat{\boldsymbol{x}}_k^l$. Note that the quantity

$$\boldsymbol{a}_{k+1}^l = \bar{\boldsymbol{x}}_{k+1}^l - \hat{\boldsymbol{x}}_{k+1}^l \tag{3.51}$$

is not necessarily zero. The sequence $\hat{\boldsymbol{x}}_k^l$ results from optimization of eqn. (6.3), which involves a tradeoff between four factors, only one of which enforces the smooth trajectory constraint defined in eqn. (3.2). Therefore, even if $\hat{\mathbf{X}}^l$ satisfies eqn. (3.2), there is no guarantee, and in fact it is unlikely that $\hat{\mathbf{X}}^{l+1}$ resulting from eqn. (3.49) will exactly satisfy eqn. (3.2).

Similarly, the actual trajectory is generated according to

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \int_{t_k}^{t_{k+1}} f(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)) d\tau. \tag{3.52}$$

For optimization, we work with a linearization around $\{\hat{\boldsymbol{x}}_k^l\}_{k=0}^K$. The model for the time propagation of the error is:

$$\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}_{k+1}^l = \boldsymbol{x}_k + \int_{t_k}^{t_{k+1}} f\big(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)\big) d\tau$$

$$-\hat{\boldsymbol{x}}_k^l - \int_{t_k}^{t_{k+1}} f\big(\bar{\boldsymbol{x}}^l(\tau), \hat{\boldsymbol{u}}(\tau)\big) d\tau.$$

$$\hat{\boldsymbol{x}}_{k+1}^l - \bar{\boldsymbol{x}}_{k+1}^l = \boldsymbol{\Phi}_k \delta \boldsymbol{x}_k^l + \big(\hat{\boldsymbol{x}}_{k+1}^l - \boldsymbol{x}_{k+1}\big) + \boldsymbol{\omega}_k$$

$$-\boldsymbol{a}_{k+1}^l = \boldsymbol{\Phi}_k \delta \boldsymbol{x}_k^l - \mathbf{I} \delta \boldsymbol{x}_{k+1}^l + \boldsymbol{\omega}_k. \tag{3.53}$$

As discussed in Section 3.1.4, the integration process that computes eqn. (3.50) also computes $\boldsymbol{\Phi}_k$ and $cov(\boldsymbol{\omega}_k) = \mathbf{Q}_k$.

3.3.4 Update Computation by Quadratic Approximation

With the notation in Section 3.3.1, the cost function in eqn. (6.4) can be rewritten as

$$\mathbf{F}(\mathbf{X}) = \|\mathbf{v}\|_{\mathbf{W}}^2 = \sum_{(i,j)\in\mathcal{M}} \|\boldsymbol{\nu}_{ij}\|_{\mathbf{R}_{ij}}^2 = \sum_{(i,j)\in\mathcal{M}} \boldsymbol{\nu}_{ij}^{\mathsf{T}} \mathbf{R}_{ij}^{-1} \boldsymbol{\nu}_{ij}. \tag{3.54}$$

Linearizing $\boldsymbol{\nu}_{ij}$ around the current estimate, it follows that

$$\boldsymbol{\nu}_{ij}(\hat{\boldsymbol{x}}_i + \delta\boldsymbol{x}_i, \hat{\boldsymbol{x}}_j + \delta\boldsymbol{x}_j) = \boldsymbol{\nu}_{ij}(\hat{\mathbf{X}} + \delta\mathbf{X})$$

$$\simeq \boldsymbol{\nu}_{ij}(\hat{\mathbf{X}}) + \mathbf{J}_{ij}\delta\mathbf{X},$$

where $\mathbf{J}_{ij} \triangleq \frac{\partial \boldsymbol{\nu}_{ij}}{\partial \mathbf{X}}\big|_{\hat{\mathbf{X}}} \in \mathbb{R}^{m_{ij} \times n(K+1)}$ is the Jacobian matrix. If rewrite

$$\mathbf{F}(\mathbf{X}) \triangleq \sum_{(i,j)\in\mathcal{M}} \mathbf{F}_{ij}(\mathbf{X}),$$

60

then each term $\mathbf{F}_{ij}(\mathbf{X})$ can be approximated as

$$\mathbf{F}_{ij}(\hat{\mathbf{X}} + \delta\mathbf{X}) = \boldsymbol{\nu}_{ij}^{\mathsf{T}}(\hat{\mathbf{X}} + \delta\mathbf{X})\mathbf{R}_{ij}^{-1}\boldsymbol{\nu}_{ij}(\hat{\mathbf{X}} + \delta\mathbf{X})$$

$$\simeq \boldsymbol{r}_{ij} + 2\boldsymbol{\eta}_{ij}^{\mathsf{T}}\delta\mathbf{X} + \delta\mathbf{X}^{\mathsf{T}}\boldsymbol{\Lambda}_{ij}\delta\mathbf{X},$$

where

$$\boldsymbol{r}_{ij} = \boldsymbol{\nu}_{ij}^{\mathsf{T}}(\hat{\mathbf{X}})\mathbf{R}_{ij}^{-1}\boldsymbol{\nu}_{ij}(\hat{\mathbf{X}}) \in \mathbb{R},$$

$$\boldsymbol{\eta}_{ij} = \mathbf{J}_{ij}^{\mathsf{T}}\mathbf{R}_{ij}^{-1}\boldsymbol{\nu}_{ij}(\hat{\mathbf{X}}) \in \mathbb{R}^{n(K+1)}, \tag{3.55}$$

$$\boldsymbol{\Lambda}_{ij} = \mathbf{J}_{ij}^{\mathsf{T}}\mathbf{R}_{ij}^{-1}\mathbf{J}_{ij} \in \mathbb{R}^{n(K+1)\times n(K+1)}. \tag{3.56}$$

Note that $\boldsymbol{\Lambda}_{ij}$ is the *information matrix* of the corresponding measurements. Thus, it follows that

$$\mathbf{F}(\hat{\mathbf{X}} + \delta\mathbf{X}) = \sum_{(i,j)\in\mathcal{M}} \mathbf{F}_{ij}(\hat{\mathbf{X}} + \delta\mathbf{X})$$

$$\simeq \sum_{(i,j)\in\mathcal{M}} \boldsymbol{r}_{ij} + 2\boldsymbol{\eta}_{ij}^{\mathsf{T}}\delta\mathbf{X} + \delta\mathbf{X}^{\mathsf{T}}\boldsymbol{\Lambda}_{ij}\delta\mathbf{X},$$

$$= \boldsymbol{r} + 2\boldsymbol{\eta}^{\mathsf{T}}\delta\mathbf{X} + \delta\mathbf{X}^{\mathsf{T}}\boldsymbol{\Lambda}\delta\mathbf{X}, \tag{3.57}$$

where

$$\boldsymbol{r} = \sum \boldsymbol{r}_{ij}, \boldsymbol{\eta} = \sum \boldsymbol{\eta}_{ij} \text{ and } \boldsymbol{\Lambda} = \sum \boldsymbol{\Lambda}_{ij}. \tag{3.58}$$

To minimize $\mathbf{F}(\hat{\mathbf{X}} + \delta\mathbf{X})$ in eqn. (3.57), the following system of equations needs to be solved

$$\boldsymbol{\Lambda}\delta\mathbf{X} = -\boldsymbol{\eta}. \tag{3.59}$$

Through *Cholesky* decomposition of the information matrix

$$\mathbf{\Lambda} = \mathbf{L}\mathbf{L}^{\mathsf{T}}, \tag{3.60}$$

$\delta\mathbf{X}$ can be derived by solving

$$\text{first } \mathbf{L}\boldsymbol{y} = -\boldsymbol{\eta} \text{ and then } \mathbf{L}^{\mathsf{T}}\delta\mathbf{X} = \boldsymbol{y},$$

where $\mathbf{L}$ is a lower-triangular matrix called the Cholesky factor of $\mathbf{\Lambda}$.

The computation of $\delta\mathbf{X}$ and the update in eqn. (3.49) will repeat until the convergence or the prespecified maximal iteration number is reached.

3.3.5 Structure of the Linearized System

The Jacobian matrix $\mathbf{J}_{ij} = \frac{\partial \boldsymbol{\nu}_{ij}}{\partial \mathbf{X}}\big|_{\hat{\mathbf{X}}}$ has the structure

$$\mathbf{J}_{ij} = \begin{bmatrix} \cdots & \mathbf{0} & \mathbf{A}_{ij} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_{ij} & \mathbf{0} & \cdots \end{bmatrix},$$

where $\mathbf{A}_{ij} \in \mathbb{R}^{m_{ij} \times n}$ is the partial of $\boldsymbol{\nu}_{ij}$ with respect to $\boldsymbol{x}_i$ and $\mathbf{B}_{ij} \in \mathbb{R}^{m_{ij} \times n}$ is the partial of $\boldsymbol{\nu}_{ij}$ with respect to $\boldsymbol{x}_j$. Note that for a standard measurement in eqn. (3.39), there is only one block in the Jacobian matrix $\mathbf{J}_{ii}$ as the partial of $\boldsymbol{\nu}_{ii}$ with respect to $\boldsymbol{x}_i$.

The corresponding information matrix $\mathbf{\Lambda}_{ij}$ in eqn. (3.56) has the structure

$$
\mathbf{\Lambda}_{ij} = \begin{bmatrix} \ddots & & & & \\ & \mathbf{A}_{ij}^{\mathsf{T}}\mathbf{R}_{ij}^{-1}\mathbf{A}_{ij} & \cdots & \mathbf{A}_{ij}^{\mathsf{T}}\mathbf{R}_{ij}^{-1}\mathbf{B}_{ij} & \\ & \vdots & & \vdots & \\ & \mathbf{B}_{ij}^{\mathsf{T}}\mathbf{R}_{ij}^{-1}\mathbf{A}_{ij} & \cdots & \mathbf{B}_{ij}^{\mathsf{T}}\mathbf{R}_{ij}^{-1}\mathbf{B}_{ij} & \\ & & & & \ddots \end{bmatrix},
$$

and $\boldsymbol{\eta}_{ij}$ has the structure

$$
\boldsymbol{\eta}_{ij} = \begin{bmatrix} \vdots \\ \mathbf{A}_{ij}^{\mathsf{T}}\mathbf{R}_{ij}^{-1}\boldsymbol{\nu}_{ij} \\ \vdots \\ \mathbf{B}_{ij}^{\mathsf{T}}\mathbf{R}_{ij}^{-1}\boldsymbol{\nu}_{ij} \\ \vdots \end{bmatrix}.
$$

Thus, in implementations $\mathbf{\Lambda}_{ij}$ and $\boldsymbol{\eta}_{ij}$ are derived by calculating the corresponding blocks, instead of evaluating the full $\mathbf{J}_{ij}$ matrices. Furthermore, $\mathbf{\Lambda}$ and $\boldsymbol{\eta}$ can be derived through summations in eqn. (3.58).

3.3.6 Computation Cost Compared with EKF

This subsection compares the computation cost of the MAP smoothing in the NRT approach with that of the EKF.
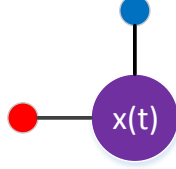
Figure 3.4: The factor graph modeling for EKF. Red factor is for the INS prior and blue factor is the aiding measurement.

The time complexity of each iteration in the optimization (3.47) is $\mathcal{O}(n^3(K+1)^3)$, since the computation is dominated by the Cholesky factorization[5] in (3.60).

The EKF update step can also be revisited from the MAP perspective (see eqn. (3.36)). Fig. shows the corresponding factor graph modeling for EKF. Furthermore, a system of equations (similar to eqn. (3.59)) can be derived by following Section 3.1.5 and Section 3.3.5,

$$\boldsymbol{\Lambda}_{kf}\delta\boldsymbol{x}_k = -\boldsymbol{\eta}_{kf}. \tag{3.61}$$

Since in each KF update eqn. (3.61) only needs be solved once, the time complexity of each KF update is approximately $\mathcal{O}(n^3)$ [72].

The computation cost of the MAP smoothing is higher than that for the EKF:

- The dimension of the information matrix $\boldsymbol{\Lambda}$ is much larger than $\boldsymbol{\Lambda}_{kf}$, since all the navigation information over the time window is considered in the optimization.

- The calculation in Section 3.3.4 may be iterated multiple times in one MAP smoothing NRT window. Multiple iterations are required when the prior error

---

[5]The complexity of the Cholesky factorization is $n^3/3$, assuming that the matrix is n-by-n.

is large relative to the nonlinear effects, to correct wrong linearization points and converge to the optimal solution.

- In each iteration, all the IMU measurements need to be reintegrated (as discussed in Section) 3.3.3 to evaluate the INS terms in eqn. (6.4) with updated estimates.

The benefit of the increased computation is enhanced accuracy and reliability.

Since the graph optimization over a time window accumulates more measurements than the filtering method, the outlier detection and identification ability of the system is enhanced. In Chapter 4, robust graph optimization method is studied.

Chapter 5 investigates RTK GPS/INS as an example of the Graph Optimization based INS.

Chapter 4

Robust Graph Optimization

Chapter 3 shows that the estimation problem in inertial navigation can be represented by probabilistic graphical models. Under standard Gaussian noise assumptions, Nonlinear Least Squares are derived for the corresponding graph optimization. This chapter is dedicated to handling outliers (i.e., faulty measurements caused by the unexpected noisy environment or the sensor failures, etc.) in (nonlinear) least squares. Several outlier handling techniques from Geodesy [4, 5, 27, 28], Computer Vision [80] and Robobtics [62] are reviewed. As a conclusion of this chapter, a hybrid approach is proposed to improve the robustness of the graph optimization associated to inertial navigation.

The following is an good example representing the problem discussed throughout this chapter,

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{n} + \boldsymbol{s}, \quad \boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}), \tag{4.1}$$

where $\boldsymbol{y} \in \mathbb{R}^m$ is the measurement vector may contains outliers, $\boldsymbol{h} : \mathbb{R}^n \mapsto \mathbb{R}^m$ is the (nonlinear) measurement model[1] of the unknown variable $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{n}$ is the unit Gaussian noise vector and $\boldsymbol{s}$ is the vector representing the errors which may make part of the measurements be outliers. In the following sections, several techniques are reviewed to estimate $\boldsymbol{x}$ when $\|\boldsymbol{s}\| \neq 0$. Furthermore, in Section 4.4 a hybrid approach is presented to improve the robustness of the graph optimization base inertial system (see Chapter 3).

Before the outlier handling techniques reviewed in the following sections, some facts based on the ANalysis Of VAriance (ANOVA) of general Least Square estimation is presented here, in order to facilitate the later analysis. Suppose $\|\boldsymbol{s}\| = 0$, and then the problem in eqn. (4.1) becomes the standard Nonlinear Least Square (NLS) problem. As reviewed in Section 3.2, NLS can be solved efficiently through iterative methods. In iterative methods, the nonlinear model is usually approximated with a linear one $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{H}\boldsymbol{x}$, where $[\boldsymbol{H}]_{m \times n}$ has full column rank. With $\|\boldsymbol{s}\| = 0$ and $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{H}\boldsymbol{x}$ the measurement model defined in eqn. (4.1) becomes

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}). \tag{4.2}$$

Since the distribution of the noise is know, the *Maximum-Likelihood* estimation could be formed to estimate $\boldsymbol{x}$,

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x} \in \mathbb{R}^n}{arg\,max}\, p_n(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}), \tag{4.3}$$

---

[1]It is required $m \geq n$ such that the problem is well-defined.

where p_n is the PDF of $\boldsymbol{n}$. Then, by evaluating the negative log-likelihood, optimization (4.3) is equivalent to

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min}\,\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2, \tag{4.4}$$

i.e., Ordinary Least Square whose closed form is well-known [37], $\hat{\boldsymbol{x}} = (\boldsymbol{H}^{\mathsf{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathsf{T}}\boldsymbol{y}$. By defining the estimation residual $\boldsymbol{r} \triangleq \boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}}$, the following proposition and corollary can be proven,

Proposition 4.1. *Letting $E[\cdot]$ to be the statistical expectation operator, it follows that*

$$E[\boldsymbol{r}^{\mathsf{T}}\boldsymbol{r}] = m - n \quad \text{and} \quad E[\boldsymbol{r}\boldsymbol{r}^{\mathsf{T}}] = \boldsymbol{I} - \boldsymbol{P}_H. \tag{4.5}$$

where $\boldsymbol{P}_H \triangleq \boldsymbol{H}(\boldsymbol{H}^{\mathsf{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathsf{T}}$.

Proof. With the definition of $\boldsymbol{r}$, it follows that

$$\boldsymbol{r} = (\boldsymbol{I} - \boldsymbol{P}_H)\,\boldsymbol{n}.$$

The matrix $\boldsymbol{P}_H$ is the projection matrix onto the range space of $\boldsymbol{H}$ (i.e., $\mathcal{C}(\boldsymbol{H})$) and $\boldsymbol{P}_H$ is symmetric, idempotent. The rank of $\boldsymbol{P}_H$ is n. For more details, check Chapter 8 of [37].

On the other hand, $\boldsymbol{I} - \boldsymbol{P}_H$ is the projection matrix onto the left nullspace of $\boldsymbol{H}$ (i.e., $\mathcal{LN}(\boldsymbol{H})$ or $\mathcal{N}(\boldsymbol{H}^{\mathsf{T}})$). So, the rank of $\boldsymbol{I} - \boldsymbol{P}_H$ is $m - n$. It can be proven that $\boldsymbol{I} - \boldsymbol{P}_H$ is also a symmetric idempotent matrix. Thus, the eigenvalues of $\boldsymbol{I} - \boldsymbol{P}_H$ are either 0 or 1. Furthermore, $E[\boldsymbol{r}\boldsymbol{r}^{\mathsf{T}}] = \boldsymbol{I} - \boldsymbol{P}_H$ is trivial since $(\boldsymbol{I} - \boldsymbol{P}_H)(\boldsymbol{I} - \boldsymbol{P}_H)^{\mathsf{T}} = \boldsymbol{I} - \boldsymbol{P}_H$.

Evaluate the expectation

$$E[\boldsymbol{r}^\mathsf{T}\boldsymbol{r}] = E[tr\{\boldsymbol{r}\boldsymbol{r}^\mathsf{T}\}]$$

$$= tr\{E[\boldsymbol{I} - \boldsymbol{P}_H]\}$$

$$= tr\{\boldsymbol{I} - \boldsymbol{P}_H\}$$

$$= m - n,$$

where $tr\{\cdot\}$ is the trace operator and the last step can be derived through the eigen-value decomposition of the real symmetric matrix $\boldsymbol{I} - \boldsymbol{P}_H$. $\qquad\square$

When the assumption about $\boldsymbol{n}$ in eqn. (4.1) replaced with $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{C})$, the measurement model can rewritten as,

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}). \tag{4.6}$$

Then, a Weighted Least Square (WLS) problem can be derived other than the OLS. The closed form solution to the WLS with measurement model (4.6) is

$$\hat{\boldsymbol{x}} = (\boldsymbol{H}^\mathsf{T}\boldsymbol{C}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^\mathsf{T}\boldsymbol{C}^{-1}\boldsymbol{y}. \tag{4.7}$$

Furthermore, the following corollary of Proposition 4.1 can be derived for WLS.

Corollary 4.2. *Given the estimated residual $\boldsymbol{r}$ from the Weighted Least Square with a weighting matrix $\boldsymbol{C}^{-1}$, it follows that*

$$E[\boldsymbol{r}^\mathsf{T}\boldsymbol{C}^{-1}\boldsymbol{r}] = m - n \quad and \quad E[\boldsymbol{r}\boldsymbol{r}^\mathsf{T}] = \boldsymbol{C} - \boldsymbol{H}(\boldsymbol{H}^\mathsf{T}\boldsymbol{C}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^\mathsf{T}. \tag{4.8}$$

Proof. The proof can be derived by defining the Cholesky decomposition of $\boldsymbol{C}^{-1} \triangleq$ $\boldsymbol{\Sigma}_c^{\mathsf{T}}\boldsymbol{\Sigma}_c$, $\bar{\boldsymbol{y}} \triangleq \boldsymbol{\Sigma}_c\boldsymbol{y}$, $\bar{\boldsymbol{H}} \triangleq \boldsymbol{\Sigma}_c\boldsymbol{H}$ and $\bar{\boldsymbol{n}} \triangleq \boldsymbol{\Sigma}_c\boldsymbol{n}$. In other words, the model $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$ is weighted by the square root matrix of the inverse covariance $\boldsymbol{C}^{-1}$ such that $\bar{\boldsymbol{n}} \triangleq$ $\boldsymbol{\Sigma}_c\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. Execute the OLS solution with the weighted measurement, then the estimated residual is

$$\bar{\boldsymbol{r}} = (\boldsymbol{I} - \boldsymbol{P}_{\bar{H}})\,\bar{\boldsymbol{n}},$$

where $\boldsymbol{P}_{\bar{H}} \triangleq \bar{\boldsymbol{H}}(\bar{\boldsymbol{H}}^{\mathsf{T}}\bar{\boldsymbol{H}})^{-1}\bar{\boldsymbol{H}}^{\mathsf{T}}$. On the other hand, $\bar{\boldsymbol{r}} = \boldsymbol{\Sigma}_c\boldsymbol{r}$. Thus, with Proposition 4.1, it follows that

$$E[\boldsymbol{r}^{\mathsf{T}}\boldsymbol{C}^{-1}\boldsymbol{r}] = E[\bar{\boldsymbol{r}}^{\mathsf{T}}\bar{\boldsymbol{r}}] = m - n, \tag{4.9}$$

$$E[\boldsymbol{r}\boldsymbol{r}^{\mathsf{T}}] = \boldsymbol{\Sigma}_c^{-1}E[\bar{\boldsymbol{r}}\bar{\boldsymbol{r}}^{\mathsf{T}}](\boldsymbol{\Sigma}_c^{\mathsf{T}})^{-1}$$

$$= \boldsymbol{\Sigma}_c^{-1}\left(\boldsymbol{I} - \boldsymbol{P}_{\bar{H}}\right)(\boldsymbol{\Sigma}_c^{\mathsf{T}})^{-1}$$

$$= (\boldsymbol{\Sigma}_c^{\mathsf{T}}\boldsymbol{\Sigma}_c)^{-1} - \boldsymbol{H}(\boldsymbol{H}^{\mathsf{T}}\boldsymbol{C}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathsf{T}} \tag{4.10}$$

and this concludes the proof. □

Note that in eqn. (4.1) and (4.2), the scalar $m - n$ is referred as the *degree-of-freedom* or DOF. Later, in Section 4.1, it is discussed that the *degree-of-freedom* of Least Square can be used to evaluate the outlier detection ability of the estimation.

4.1 Receiver Autonomous Integrity Monitoring

Receiver Autonomous Integrity Monitoring (RAIM) techniques are widely discussed in Geodesy community (see [27, 28] and references). RAIM methods are based on Baarda's seminal works of hypothesis test procedures used in Geodesy and Surveying [4]. Conventional RAIM is for GNSS-only [28]. In [27], Extended RAIM (eRAIM) was proposed for GPS/INS with Kalman Filtering (see Section 3.1.5) which incorporates the current GPS measurement and the one step prior. The classical RAIM method is comprised of the 'Detection-Identificationi-Adaption (DIA) trilogy'.

To facilitate the later discussion of the eRAIM application on EKF based GPS/INS (Section 4.4.1), the Weighted Least Square problem (4.6) is used in this section as an example to show the basic idea of the RAIM. The method can be applied to OLS straightforwardly by setting the weighting matrix $\boldsymbol{C} = \boldsymbol{I}$.

Section 4.4 extends the RAIM framework to the graph optimization based INS detailed in Chapter 3.

4.1.1 Detection

Let $\boldsymbol{r}(\hat{\boldsymbol{x}})$ be the estimated residuals from the ordinary least square contains outlier measurements. With the Corollary 4.2, it follows that

$$E[\boldsymbol{r}(\hat{\boldsymbol{x}})^\intercal \boldsymbol{C}^{-1} \boldsymbol{r}(\hat{\boldsymbol{x}})] = m - n \quad \text{and} \quad E[\boldsymbol{r}(\hat{\boldsymbol{x}}) \boldsymbol{r}(\hat{\boldsymbol{x}})^\intercal] = \boldsymbol{C} - \boldsymbol{H}(\boldsymbol{H}^\intercal \boldsymbol{C}^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^\intercal.$$

Then, the *a-posteriori* variance factor test [28] is introduced to check validity of the model in eqn. (4.2) and the capability to detect outliers. The variance factor is determined as

$$\hat{\boldsymbol{\sigma}}_0^2 = \frac{\boldsymbol{r}^\intercal(\hat{\boldsymbol{x}})\boldsymbol{C}^{-1}\boldsymbol{r}(\hat{\boldsymbol{x}})}{m-n}. \tag{4.11}$$

The detection procedure is concluded by testing the variance factor against the two-tailed limits derived from Chi-squared distribution

$$\frac{\boldsymbol{\chi}_{\alpha/2,m-n}^2}{m-n} < \hat{\boldsymbol{\sigma}}_0^2 < \frac{\boldsymbol{\chi}_{1-\alpha/2,m-n}^2}{m-n}, \tag{4.12}$$

where $m-n$ is the DOF of the Chi-squared distribution and the α is the significant level of the test. Often times, α is picked as 0.05.

If the test fails, then the model in eqn. (4.2) is considered as invalid. The invalidity can be caused by many reasons, e.g. modeling error, outliers and poor characterization of sensors (poor $\boldsymbol{C}$ assumption). Herein, we assume only outliers can cause the test failure.

4.1.2 Identification

For the case that the i-th measurement has outliers, the model in eqn. (4.2) has to be extended to the form [24, 28, 38],

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}_i s_i + \boldsymbol{n}, \quad \boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{C}), \tag{4.13}$$

where $\boldsymbol{e}_i = [0,\ldots,0,1,0,\ldots,0]^\intercal \in \mathbb{R}^m$ (only the i-th element is 1) and s_i is the outlier.

The identification of an outlier depends on the statistical testing where the null hypothesis is that the measurements are outlier free and the alternative hypothesis is that an outlier exists:

$$\text{Null Hypothesis} \quad : \quad E(\hat{s}_i) = 0; \qquad (4.14)$$

$$\text{Alternative Hypothesis} \quad : \quad E(\hat{s}_i) = s_i \neq 0. \qquad (4.15)$$

In identification, the w-test [4, 38, 67] is executed to test each measurement. For the i-th measurement of $\boldsymbol{y}$, the test statistics is

$$w_i = \frac{\boldsymbol{e}_i^\mathsf{T} \boldsymbol{C}^{-1} \boldsymbol{r}(\hat{\boldsymbol{x}})}{\sqrt{\boldsymbol{e}_i^\mathsf{T} \boldsymbol{C}^{-1} \boldsymbol{Q} \boldsymbol{C}^{-1} \boldsymbol{e}_i}}, \qquad (4.16)$$

where $\boldsymbol{Q} \triangleq E[\boldsymbol{r}(\hat{\boldsymbol{x}})\boldsymbol{r}(\hat{\boldsymbol{x}})^\mathsf{T}] = \boldsymbol{C} - \boldsymbol{H}(\boldsymbol{H}^\mathsf{T} \boldsymbol{C}^{-1} \boldsymbol{H})^{-1} \boldsymbol{H}^\mathsf{T}$ is the *posteriori* covariance of the estimated residuals $\boldsymbol{r}(\hat{\boldsymbol{x}})$. Under the null hypothesis, $\boldsymbol{w}_i$ has a standard normal distribution; for a outlier s_i, w_i has the following noncentrality,

$$\boldsymbol{\delta}_i = s_i \sqrt{\boldsymbol{e}_i^\mathsf{T} \boldsymbol{C}^{-1} \boldsymbol{Q} \boldsymbol{C}^{-1} \boldsymbol{e}_i}.$$

Thus, the statistics w_i is tested against normality tests with certain confidence level $1 - \alpha$, e.g.,

$$\mathcal{N}^{-1}(\alpha/2, 0, 1) < w_i < \mathcal{N}^{-1}(1 - \alpha/2, 0, 1), \qquad (4.17)$$

where $\mathcal{N}^{-1}(x, 0, 1)$ is the inverse normal distribution function. Note that if $\boldsymbol{C} = \boldsymbol{I}$ (i.e., the measurements are already normalized), the w-test statistics becomes $w_i = \boldsymbol{r}_i(\hat{\boldsymbol{x}})/\sqrt{\boldsymbol{e}_i^\mathsf{T} \boldsymbol{Q} \boldsymbol{e}_i}$, and there is a threshold (a.k.a. critical value in hypothesis test

terminology) $\tau_i > 0$ can be found as

$$\tau_i = \sqrt{\boldsymbol{e}_i^\intercal \boldsymbol{Q} \boldsymbol{e}_i} \cdot \mathcal{N}^{-1}(1 - \alpha/2, 0, 1), \tag{4.18}$$

such that if $|\boldsymbol{r}_i(\hat{\boldsymbol{x}})| > \tau_i$, the alternative hypothesis (4.14) is accepted. This w-test is executed multiple times until no outlier identified.

4.1.3 Adaptation

Once the identification procedure goes over all the measurements, the one with the largest residual (as a single outlier may cause multiple test failures [29]) will be eliminated from the estimation. After removing the identified outlier, the ordinary least square estimation is executed again, and then the detection and identification procedures repeat. If another outlier is found it is removed from the model and the measurement that was first regarded as an outlier is reincluded and the model retested. This procedure repeats until no more outliers can be identified.

4.2 Cost Regularization Approach

Besides the hypothesis test way to identify and remove the outliers in the measurements, regularization methods, e.g. l_1 regularization, are more popular in Machine Learning and Computer Vision research community [26, 73]. Regularization methods aim to modify the original least square cost (i.e., 2-norm) by different way to

realize robustness. This section presents the Huber loss function and the Least Soft-Thresholding Square (LSS) method.

4.2.1 Huber Loss Function

To make the standard Least Square formulation robust to outliers in measurements, robust cost terms can be included to regularize the original one. This idea is from the observation that the l_2-norm $\| \cdot \|_2^2$ used by the standard least square increases sharply when the magnitude of the residual goes up. The purpose of adding robust terms is to less-weighting the large residuals which is more likely from a outlier measurements. The l_1-norm $\| \cdot \|_1$ can achieve robustness but it is hard to solve a l_1-norm minimization since the absolute value function is undifferentiable. Instead, Huber's loss [31, 32] or Tukey's biweight [30] is often used. The Huber's loss function is defined as follows,

$$L_u(x) = \begin{cases} x^2, & |x| \leq \lambda; \\ 2\lambda|x| - \lambda^2, & |x| > \lambda. \end{cases} \tag{4.19}$$

Fig. 4.1 presents the l_1, l_2 and Huber cost over the interval $[-3, 3]$.

It can be seen that the Huber loss function weights the large residuals less than the l_2 norm, but more than the l_1 norm. Later in Section 4.2.2, it will be shown that Huber loss function actually combines the l_2 norm and the l_1 norm. In other words, Huber loss function regularize the original l_2 cost with a l_1 norm term. The threshold λ indicates the strength of the regularization. By replacing the l_2 norm in the OLS
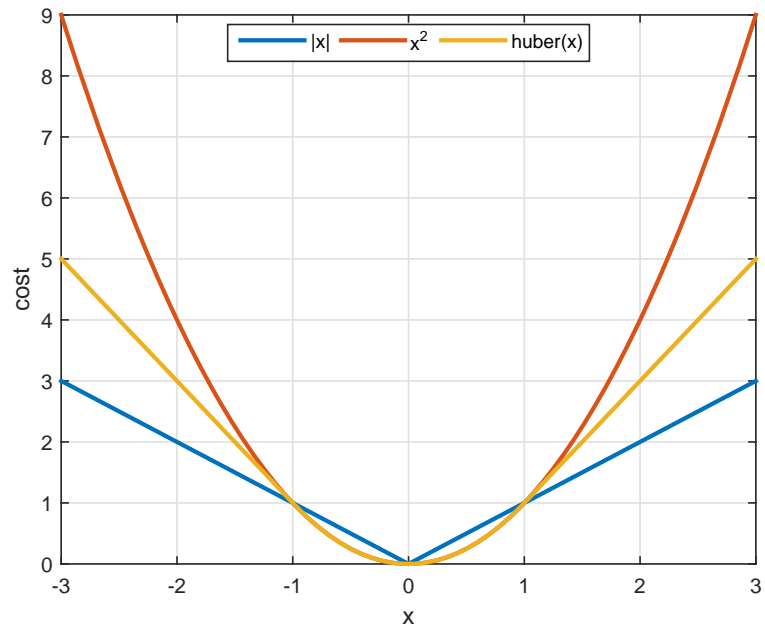
Figure 4.1: l_1 norm, l_2 norm and Huber cost.

with the Huber loss, the optimization (4.4) is robustified as

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x} \in \mathbb{R}^n}{arg\,min} \sum_i L_u(y_i - \boldsymbol{h}_i \boldsymbol{x}), \tag{4.20}$$

where $y_i \in \mathbb{R}$ is i-th measurement and $\boldsymbol{h}_i \in \mathbb{R}^{1 \times n}$ is the i-th row of $\boldsymbol{H}$. Furthermore, by considering the identical KKT condition [6], it can be proven that the optimization in (4.20) is equivalent to

$$\min_{(\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{w} \in \mathbb{R}^m)} \sum_i \frac{(y_i - \boldsymbol{h}_i \boldsymbol{x})^2}{w_i + 1} + \lambda^2 \sum_i w_i, \text{ s.t. }, w_i \geq 0, \tag{4.21}$$

where $\boldsymbol{w}$ is a variable weight vector. Thus, one classical method to solve the Huber-robust least square in eqn. (4.20) is *iteratively reweighted least square* (IRLS).

In the graph optimization library $g2o$ [41], Huber loss cost is used for robust least square.

4.2.2 Least Soft-Thresholding Square (LSS)

In [80], a Least Soft-Thresholding Square (LSS) approach is presented to provide more efficient way to solve eqn. (4.20).

In LSS framework, the outlier term $\boldsymbol{s}$ in eqn. (4.1) is modeled as an $i.i.d$ Laplacian noise vector, i.e., $s_i \in \mathcal{L}(0, \sigma_L)$. Then, when a linear model is considered, a *Maximum-*

Likelihood estimation is formed through maximizing

$$p(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{s})$$

$$= p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{s})p(\boldsymbol{x}, \boldsymbol{s}) \tag{4.22}$$

$$= p_n(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x} - \boldsymbol{s})p_L(\boldsymbol{s})$$

$$= Kexp\left\{ -\frac{1}{2}\left(\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x} - \boldsymbol{s}\|_2^2 + \lambda\|\boldsymbol{s}\|_1 \right) \right\},$$

where $K = (\sqrt{2}\sigma_L)^{-m}(\sqrt{2\pi})^{-m}$ and $\lambda = 2\sqrt{2}/\sigma_L$. By taking negative log-likelihood of eqn. 4.22 and get rid of the constant coefficient, the following cost function can be get as a Least Soft-threshold Square problem

$$L_s(\boldsymbol{x}, \boldsymbol{s}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x} - \boldsymbol{s}\|^2 + \lambda\|\boldsymbol{s}\|_1. \tag{4.23}$$

Note that the λ in $L_s(\boldsymbol{x}, \boldsymbol{s})$ also indicates the strength of the l_1 regularization. In fact, it can be proven that the cost formed by the Huber loss function in eqn. (4.20) is equivalent that of LSS in eqn. (4.23), if the λ is identical.

To minimize the LSS in 4.23, the following two useful insights are presented in [80],

- Given $\hat{\boldsymbol{s}}$, the optimal $\hat{\boldsymbol{x}}$ can be obtained by the OLS solution

$$\hat{\boldsymbol{x}} = (\boldsymbol{H}^\mathsf{T}\boldsymbol{H})^{-1}\boldsymbol{H}^\mathsf{T}(\boldsymbol{y} - \hat{\boldsymbol{s}}).$$

- Given $\hat{\boldsymbol{x}}$, the optimal $\hat{\boldsymbol{s}}$ can be obtained by a soft-thresholding (or shrinkage [73]) operation

$$\hat{s}_i = \mathcal{S}_\lambda(|y_i - \boldsymbol{h}_i\boldsymbol{x}|),$$

78

where

$$\mathcal{S}_\lambda(x) = \max(|x| - \lambda, 0) sgn(x)$$

and $sgn(\cdot)$ is the sign function.

The following is the algorithm to solve LSS.

Algorithm 4.1: Least Soft-Thresholding Square (LSS)

> **Input** : An measurement vector $\boldsymbol{y}$, matrix $\boldsymbol{H}$, pre-computed matrix
>
> $\boldsymbol{P} = (\boldsymbol{H}^\intercal \boldsymbol{H})^{-1} \boldsymbol{H}^\intercal$ and a constant λ

1 Initialize iteration index $l = 0$ and $\boldsymbol{s}^0 = \boldsymbol{0}$;

2 **while** *not converged and not terminated* **do**

3   Obtain $\boldsymbol{x}^{l+1}$ via $\boldsymbol{x}^{l+1} = \boldsymbol{P}(\boldsymbol{y} - \boldsymbol{s}^l)$;

4   Obtain $\boldsymbol{s}^{l+1}$ via $\boldsymbol{s}^{l+1} = \mathcal{S}_\lambda(|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}^l|)$;

5 $l \leftarrow l + 1$

6 **end**

> **Output**: $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{s}}$

Similar algorithm for WLS problem can also be derived straightforwardly by plugging in the weighting matrix $\boldsymbol{C}^{-1}$.

An useful insight of the Huber loss function method and the LSS method is on the different ways they deal with the outliers:

- Huber loss function method aims to put less weight on the outliers.

- LSS method aims to compensate the outliers.

The advantage of the LSS method over the standard Huber loss function is that in LSS the matrix $\boldsymbol{P}$ needs evaluating once before the iterations. However, due to the reweighted mechanism in the optimization with Huber loss function, the matrix $\boldsymbol{P}$ needs re-evaluation for every iteration as

$$\boldsymbol{P} = (\boldsymbol{H}^{\mathsf{T}}(\boldsymbol{W}^l)^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathsf{T}}(\boldsymbol{W}^l)^{-1}, \tag{4.24}$$

where $\boldsymbol{W}_i$ is the weighting matrix adjusted by the IRLS in the i-th iteration.

4.3 Hypothesis Test Aided LSS (HT-LSS)

Although the applications of LSS show good performance and efficiency, picking the right λ is still heuristic in the literature. For example, in [80] the λ is required as a "small" constant and picked as $\lambda = 0.20$ for all the experiments. No theoretical derivation is given for this choice.

This section explores a way to set λ based on the hypothesis tests used in RAIM (see Section 4.1). The outlier handling strategies of RAIM and cost regularization methods (see Section 4.2) can be compared as:

- After the detection of the existence of outliers, RAIM aims to identify the outlier measurements and remove them from the estimation formulation. The identification is realized through hypothesis tests (e.g. w-test) of the estimated residual against some distribution (e.g. normal distribution) with certain confidence level (e.g. $1 - \alpha = 0.95$).

- Cost regularization aims to identify the outliers, and to reduce the outliers'
  effect on the estimation result through making them conform with the inliers
  (e.g. re-weighting, shrinkage). The identification is realized by checking the
  estimated residual with some prespecified threshold λ.

The above comparison indicates that in both RAIM and cost regularization method,
there is a threshold picked to judge whether an estimated residual belongs to an out-
lier measurement or not. The difference is that in RAIM this threshold (i.e., τ_i) is
picked rigorously from proper hypothesis test with a specified confidence level and
a nominated distribution. So, the threshold λ in LSS can also be set based on the
critical value in the hypothesis tests of RAIM.

Algorithm 3.2 shows the procedure of the proposed Hypothesis-Test based Least
Soft-threshold Square (HT-LSS) which can also outlined as Detection-Identification-
Adaption (DIA). Note that in Algorithm 3.2 the threshold λ is set based on the
hypothesis test discussed in Section 4.1. Furthermore, in Algorithm 3.2, a one-tail
version of the *a-posteriori* variance factor test (4.27) is applied before starting the
shrinkage operations. This variance factor test aims to detect the existence of the
outliers and may save computation versus the original LSS algorithm if there is no
outliers.

Algorithm 4.2: Hypothesis Test based Least Soft-Thresholding Square

    Input : An measurement vector $\boldsymbol{y}$, matrix $\boldsymbol{H}$, pre-computed matrix

$$\boldsymbol{P} = (\boldsymbol{H}^\intercal \boldsymbol{H})^{-1} \boldsymbol{H}^\intercal \text{ and } \boldsymbol{Q} = \boldsymbol{I} - \boldsymbol{H}\boldsymbol{P}, \text{ a confidence level } 1 - \alpha.$$

1 for $i = 1{:}m$ **do**

2   Set $\lambda_i \leftarrow \tau_i = \sqrt{\boldsymbol{e}_i^\intercal \boldsymbol{Q} \boldsymbol{e}_i} \cdot \mathcal{N}^{-1}(1 - \alpha/2, 0, 1)$;

3 end

4 Initialize $l = 0$, $\boldsymbol{s}^0 = \boldsymbol{0}$, $\boldsymbol{x}^0 = \boldsymbol{P}\boldsymbol{y}$, $\boldsymbol{r}^0 = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}^0$;

5 while *not converged or* $\|\boldsymbol{r}^l\|^2 > \chi^2_{\alpha, m-n}$ **do**

6   **for** $i = 1{:}m$ **do**

7   Obtain s_i^{l+1} via $s_i^{l+1} = \mathcal{S}_{\lambda_i}(|y_i - \boldsymbol{h}_i \boldsymbol{x}^{l+1}|)$;

8   **end**

9   Obtain $\boldsymbol{x}^{l+1}$ via $\boldsymbol{x}^{l+1} = \boldsymbol{P}(\boldsymbol{y} - \boldsymbol{s}^{l+1})$;

10   Obtain $\boldsymbol{r}^{l+1}$ via $\boldsymbol{r}^{l+1} = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}^{l+1}$;

11 $l \leftarrow l + 1$

12 end

    Output: $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{s}}$

The advantage of the proposed HT-LSS approach over the RAIM is that the tedious remove-and-reinclude book-keeping in the adaption procedure is avoid. Furthermore, since LSS is actually to model the error term $\boldsymbol{s}$, for GNSS applications multipath error modeling could be incorporated to improve the outlier handling performance. In RAIM, identified outlier measurements are just abandoned, without

considering their correlations. The advantage of the proposed HT-LSS approach over the Huber loss function method is that in HT-LSS the matrix $\boldsymbol{P}$ only needs evaluating once before the iterations, versus the iterative reweighting procedure for Huber loss function in eqn. (4.24). The advantage of the proposed HT-LSS approach over the original LSS with heuristic threshold λ is that λ is set from the critical value of a proper hypothesis test with certain confidence level. Original LSS with heuristic threshold λ may identify more or less outliers than expected.

In the rest of this section, a simulation example is given to show the performance of the proposed method. In the simulation, measurements are made through the following model,

$$\tilde{y} = x + 1 + s + n, \quad n \sim \mathcal{N}(0, 1),$$

where s is the injected outliers. To fitting this linear model $y = ax + b$, totally 122 measurements are made randomly over the interval $x \in [0, 10]$. So, the degree-of-freedom of this problem is 120. This DOF matches the situation of the smoothing based dual frequency GPS/INS over a 10-second window with 7 satellites available. For evaluation purpose, 10% of the measurements are randomly picked and injected outliers with random magnitude S from -9.7 to 10.3, see Table 4.3:

Table 4.3 concludes the fitting results with different methods discussed above. For this case, the proposed HT-LSS is the best method.

Fig. 4.2 shows the OLS result with the contaminated data. The right-up plot of Fig. 4.2 is the histogram of estimated residuals and several large residuals can be

Table 4.1: Outlier Injected List

Meas. Index	Outlier
2	7.122
4	-1.939
14	3.096
17	0.559
21	1.444
27	3.621
47	0.020
68	1.955
88	-1.449
99	0.273
108	3.571
113	7.470

Table 4.2: Line fitting with outlier measurements

Method	Result
Truth	$(1.0000, 1.0000)$
OLS	$(1.2382, 0.9964)$
RAIM	$(1.1604, 1.0018)$
LSS($\lambda = 0.2$)	$(1.1993, 0.9917)$
HT-LSS	$(1.0993, 1.0032)$

observed over 5. In the right-down plot of Fig. 4.2, yellow cross marks the residual corresponding to the measurement injected outliers. Fig. 4.3 shows the RAIM result. Note that the number of residuals is smaller than 120 since multiple measurements are identified as outliers and removed. Fig. 4.4 shows the result of the original LSS method with $\lambda = 0.2$. The right-up plot of Fig. 4.4 shows that LSS treats any measurement (even an inlier) with residual over 0.2 as an outlier and 'shrinks' its effect. Due to this tight shrinkage, there are two peaks at ± 0.20 in the right-up plot of Fig. 4.4, and the two boundaries aligned with ± 0.20 in the right-down plot of Fig. 4.4. The original LSS with $\lambda = 0.2$ identifies much more measurements as outliers than the true amount, and this may impair the fitting performance when the DOF is low. Fig. 4.5 shows the performance of the proposed HT-LSS which provide the most accurate fitting in this test.

In Section 4.4.2, the proposed HT-LSS approach is applied in the graph optimization based INS which is presented in Chapter 3.

4.4 Robust GOINS

This section proposes a novel method of outlier-rejection/robust-optimization in inertial navigation. First, the Extended RAIM (eRAIM) approach from Geodetic community is reviewed. Then, a novel approach is proposed by extending the HT-
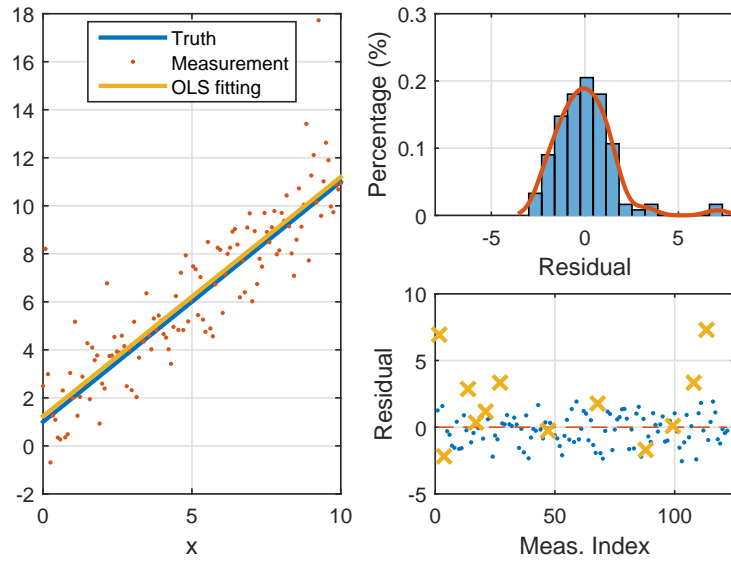
Figure 4.2: OLS with outlier measurements. The right-up plot shows the histogram of the estimated residuals. In the right-down plot, yellow cross marks the residual corresponding to the measurement injected outliers.
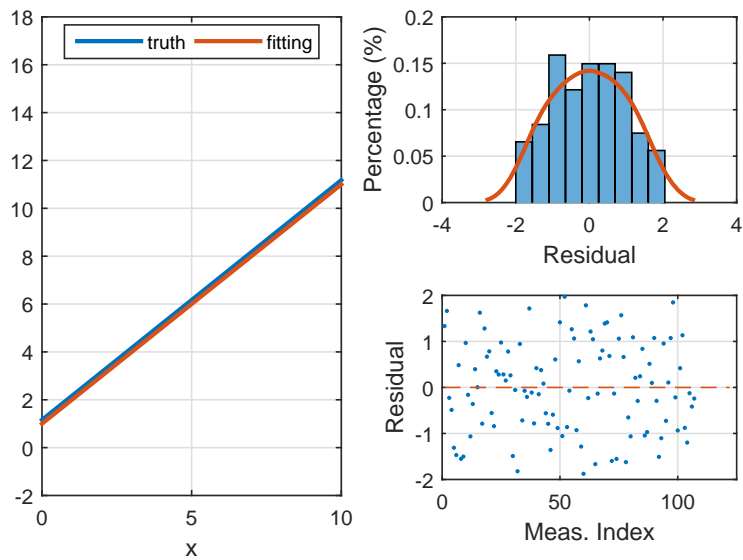
Figure 4.3: RAIM result.
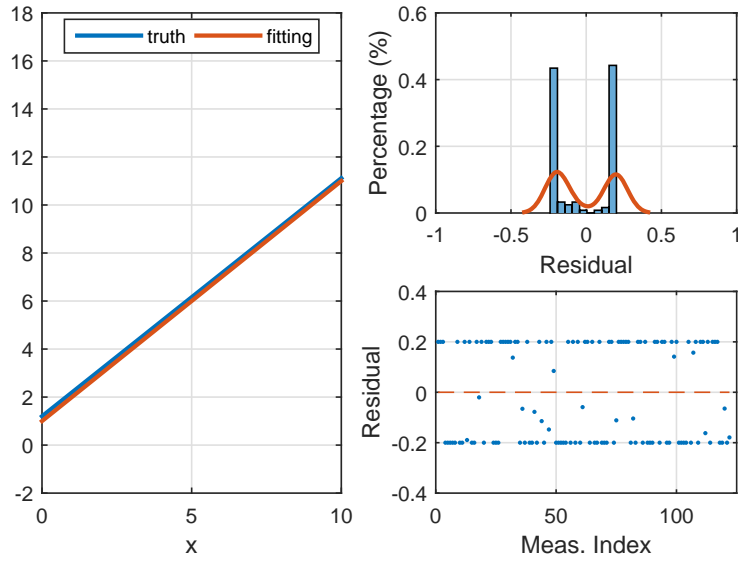
Figure 4.4: Original LSS method with $\lambda = 0.2$.
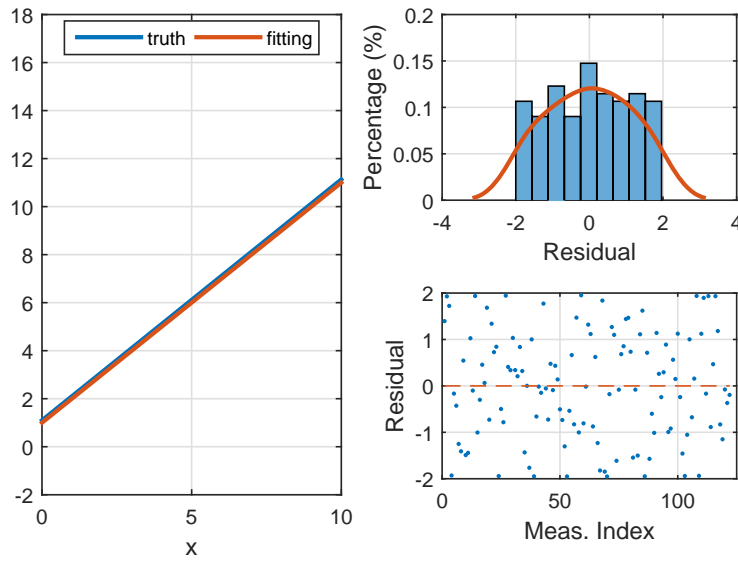
Figure 4.5: Hypothesis Test based LSS.

LSS proposed in Section 4.4.2 to the graph optimization based INS presented in Chapter 3.

4.4.1   eRAIM for EKF based GPS/INS

In Section 3.1.5, the EKF for INS is revisited from the Weighted Least Square (WLS) perspective. Herein the analysis in Section 3.1.5 is extend, and in this case the posteriori residual is defined as

$$\boldsymbol{r}(\delta \boldsymbol{x}_k^+) = \delta \boldsymbol{Z}_k - \mathbf{A}_k \delta \boldsymbol{x}_k^+$$

With the Corollary 4.2, it follows that

$$E[\boldsymbol{r}(\delta \boldsymbol{x}_k^+)^\intercal \boldsymbol{C}_k^{-1} \boldsymbol{r}(\delta \boldsymbol{x}_k^+)] = n_m \quad \text{and} \quad E[\boldsymbol{r}(\delta \boldsymbol{x}_k^+)\boldsymbol{r}(\delta \boldsymbol{x}_k^+)^\intercal] = \boldsymbol{C}_k^{-1} - \mathbf{A}_k \mathbf{P}_k^+ \mathbf{A}_k^\intercal, \quad (4.25)$$

where n_m is the number of aiding measurements at t_k and $\mathbf{P}_k^+ = \left(\mathbf{A}_k^\intercal \boldsymbol{C}_k^{-1} \mathbf{A}_k\right)^{-1}$ is the covariance of $\delta \boldsymbol{x}_k^+$.

Note that in eqn. (4.25), the *degree-of-freedom* n_m is derived from $(n_s + n_m) - n_s$, i.e., the difference between the total number of measurements (INS prior and aiding ones) and that of unknown variables (the current INS state) to be estimated. The corresponding variance factor is determined as

$$\hat{\boldsymbol{\sigma}}_k^2 = \frac{\boldsymbol{r}(\delta \boldsymbol{x}_k^+)^\intercal \boldsymbol{C}_k^{-1} \boldsymbol{r}(\delta \boldsymbol{x}_k^+)}{n_m}. \tag{4.26}$$

The detection procedure is concluded by testing the variance factor against the two-tailed limits derived from Chi-squared distribution

$$\frac{\boldsymbol{\chi}^2_{\alpha/2,n_m}}{n_m} < \hat{\boldsymbol{\sigma}}^2_k < \frac{\boldsymbol{\chi}^2_{1-\alpha/2,n_m}}{n_m}, \tag{4.27}$$

where n_m is the DOF of the Chi-squared distribution and the α is the significant level of the test. Often times, α is picked as 0.05.

If the test fails, then the model in eqn. (3.33-3.34) is considered as invalid. The invalidity can be caused by many reasons, e.g. modeling error, outliers and poor characterization of sensors (poor $\boldsymbol{C}_k$ assumption). Herein, we assume only outliers can cause the test failure.

For the case that the i-th measurement has outliers, the linearized measurement model (see Section 3.1.5)

$$\delta \boldsymbol{Z}_k = \mathbf{A}_k \delta \boldsymbol{x} + \boldsymbol{v}_k, \quad \boldsymbol{v}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{C}_k) \tag{4.28}$$

has to be extended to the form [24, 28, 38],

$$\delta \boldsymbol{Z}_k = \mathbf{A}_k \delta \boldsymbol{x} + \boldsymbol{e}_i s_i + \boldsymbol{v}_k, \quad \boldsymbol{v}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{C}_k), \tag{4.29}$$

where $\boldsymbol{e}_i = [0, \ldots, 0, 1, 0, \ldots, 0]^\intercal$ (only the i-th element is 1) and s_i is the outlier.

The identification of an outlier depends on the statistical testing where the null hypothesis is that the measurements are outlier free and the alternative hypothesis is

that an outlier exists:

$$\text{Null Hypothesis} \quad : \quad E(\hat{s}_i) = 0;$$

$$\text{Alternative Hypothesis} \quad : \quad E(\hat{s}_i) = s_i \neq 0.$$

In identification, the w-test [4, 38, 67] is executed to test each measurement. For the i-th measurement $\delta \boldsymbol{Z}_k$, the test statistics is

$$w_i = \frac{\boldsymbol{e}_i^\intercal \boldsymbol{C}_k^{-1} \boldsymbol{r}(\delta \boldsymbol{x}_k^+)}{\sqrt{\boldsymbol{e}_i^\intercal \boldsymbol{C}_k^{-1} \boldsymbol{Q}_{r_k} \boldsymbol{C}_k^{-1} \boldsymbol{e}_i}}. \tag{4.30}$$

Under the null hypothesis, $\boldsymbol{w}_i$ has a standard normal distribution; for a outlier s_i, w_i has the following noncentrality,

$$\boldsymbol{\delta}_i = s_i \sqrt{\boldsymbol{e}_i^\intercal \boldsymbol{C}_k^{-1} \boldsymbol{Q}_{r_k} \boldsymbol{C}_k^{-1} \boldsymbol{e}_i}.$$

Thus, the statistics w_i is tested against normality tests, e.g. *Kolmogorov-Smirnov* test. This w-test is executed multiple times until no outlier identified.

Once the outlier is identified, it will be removed from the measurement set. After removal the identified outliers, the estimation is done again, and then the detection and identification procedures repeat until no detection reported.

This eRAIM approach was reported efficient for EKF based GPS/INS [27]. As stated in RAIM [28], the ability of estimation to detect and handling the outliers depends on the degree-of-freedom of problem. However, in outlier-prone situations, the availability of the GPS satellites is also limited, e.g. in urban canyon. As discussed

in Chapter 3, the graph optimization based approach could enhance the degree-of-freedom since the smoothing window may accumulate more measurements. Thus, the graph optimization based INS maintains better robustness to outlier measurements. Section 4.4.2 proposes a novel method and shows an example for GPS/INS.

4.4.2 Robust GOINS based on HT-LSS

In this section, the HT-LSS method proposed in Section 4.4.2 is extended to the graph optimization based INS presented in Chapter 3. The GPS/INS application is focused on and the discussed method can be extended to other INS implementations straightforwardly.

As discussed in Section 4.1 and its references, the *degree-of-freedom* of an estimation problem determines the ability to detect outliers and the estimation robustness. In the graph optimization based INS presented in Chapter 3, a smoothing problem over a time window is considered for each navigation state estimation. Since more measurements are included over a window than just the measurements of a single epoch, the graph optimization based INS over a time window can improve the navigation system's ability to detect and handling outliers. Considering the problem of the graph optimization based INS as Nonlinear Least Square, the HT-LSS method proposed in Section 4.4.2 can be extended as follows:

1. (Detection) After the convergence of the NLS in eqn. (3.47), do the following one-tail[2] *variance factor* test,

$$\|\boldsymbol{r}(\hat{\mathbf{X}})\|^2 < \chi^2_{1-\alpha/2, M},$$

where M is the total number of aiding measurements and the degree-of-freedom $dof \triangleq [n_s(K+1) + M] - n_s(K+1) = M$ is the difference between the total number of unknowns (i.e., $\mathbf{X} \in \mathbb{R}^{n_s(K+1)}$) and the total number of measurements (including prior distribution n_s, INS constraint $n_s K$, GPS measurements M). The failure of the test implies the existence of the outlier, and the following outlier handling procedure proceed. Otherwise, the robust optimization is done and the final navigation estimates can be published.

2. It is assumed that there is no outlier in INS constraints, and only outliers in prior constraints and aiding measurements are considered. Model the estimated residuals of these measurements as

$$\boldsymbol{r} = \boldsymbol{\Sigma}_{\mathbf{R}} \mathbf{J} \delta \mathbf{X} + \boldsymbol{s} + \boldsymbol{n}, \quad \boldsymbol{n} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}),$$

where $\boldsymbol{\Sigma}_{\mathbf{R}}$ is the normalizing matrix defined in Section 3.2, $\mathbf{J}$ is the Jacobian matrix of the corresponding measurements, $\delta \mathbf{X}$ is the trajectory update from the last iteration of NLS solution. Then, the HT-LSS shrinkage is applied on $\boldsymbol{s}$

$$\boldsymbol{s} = \mathcal{S}_{\boldsymbol{\tau}}(\boldsymbol{r} - \boldsymbol{\Sigma}_{\mathbf{R}} \mathbf{J} \delta \mathbf{X}),$$

---

[2]In practice, we usually do not care about the situation where the cost is too small.

where $\boldsymbol{\tau}$ is a vector of the critical values derived from the w-test (4.18) for each weighted measurement.

3. Solve the NLS again through evaluating the residuals as $\boldsymbol{r}' = \boldsymbol{r} - \boldsymbol{s}$. Repeat these three steps until no failures in the *variance factor* test or $\hat{\mathbf{X}}$ and $\boldsymbol{s}$ converged.

One advantage of the proposed method for GPS/INS application is that the Jacobian matrix $\mathbf{J}$ for GPS measurements does not need re-evaluating multiple times after Step (1). This is due to the weak nonlinearity and slow time-varying property of GPS measurement model. These special properties are investigated mathematically and utilized creatively in Chapter 6. In contrast, if Huber loss function method is applied here, the weighted Jacobian matrix needs to be re-calculated after the re-weighting $\mathbf{R} \leftarrow \mathbf{W}^l \mathbf{R}$.

In the rest of this section, an example of the graph optimization based GPS/INS is presented to illustrate the proposed HT-LSS based robust navigation. In this experiment, a 10-second IMU plus GPS data is picked. In this data, the L1/L2 pseudorange measurements are inspected manually with the RTK GPS/INS smoothing result, to ensure there is no outliers. For performance validation purpose, epochs and satellites are randomly picked to inject outliers as in Table 4.4.2. Thus, in this experiment we have three data sets to work on:

1. Original data without any artificial modification.

2. Outlier Injected data which is injected outliers as in Table 4.4.2.

3. Outlier Cleaned data in which the measurements with injected outliers are removed.

Then, with these three data sets, four results can be obtained:

1. Smoothing (without outlier handling) with Data (1).

2. Smoothing (without outlier handling) with Data (2).

3. Smoothing (with outlier handling) with Data (2).

4. Smoothing (without outlier handling) with Data (3).

An useful insight on the four tests is that once a measurement is contaminated with an outlier error, no useful information for estimation can be recovered from it anymore. The best option is to remove it from the formulation. Thus, one the data is contaminated with outliers, the result of Test (1) can never be recovered. The best result we can expected is one close to that of Test (4).

Table 4.3: Outlier Injected List

Time Index	PRN	L1 Outlier (m)	L2 Outlier (m)	L1 Identified	L2 Identified
2	1	4.8697	2.5125	√	√
3	13	-2.5746	1.9839	√	√
3	32	4.4479	-1.8058	√	√
4	23	0.5293	-0.9808		√
6	32	0.7515	-2.4864		√
9	16	2.5903	3.8155	√	√

Fig. 4.6 compares the trajectories from the four tests on tangent plane. It can be observed that the results of Test (3) and (4) are close to each other, and both of them are more close to the Test (1) result which can be considered as ground truth, than that of Test (2). The result of Test (2) is the most deviated from that of Test (1). Fig. 4.7 and 4.8 shows the 3D positioning errors between the results of Test (2-4) and Test (1). Fig. 4.9 shows the positioning errors between the results of Test (2-3) and Test (4). It can be observed that the result of Test (3) is much closer to that of Test (4) than Test (2). It implies that the proposed HT-LSS approach identifies and handles the outliers properly. In the last two columns of Table 4.4.2, the contaminated measurements are marked if identified by the w-test as an outlier. There are two contaminated measurements are not identified since the injected errors are relatively small such that the resulted residuals are not over the critical values. The histograms of the estimated L1/L2 pseudorange residuals (compensated with $\boldsymbol{s}$, i.e. $\boldsymbol{r} - \boldsymbol{s}$) in Fig. 4.10 - 4.12 for Test (2-4) also illustrate this point. Fig. 4.13 shows the estimated L1/L2 pseudorange residuals (uncompensated, i.e., $\boldsymbol{r}$) and $\boldsymbol{s}$. By this way, $\boldsymbol{s}$ can be regarded as estimates of the errors in the weighted measurements.
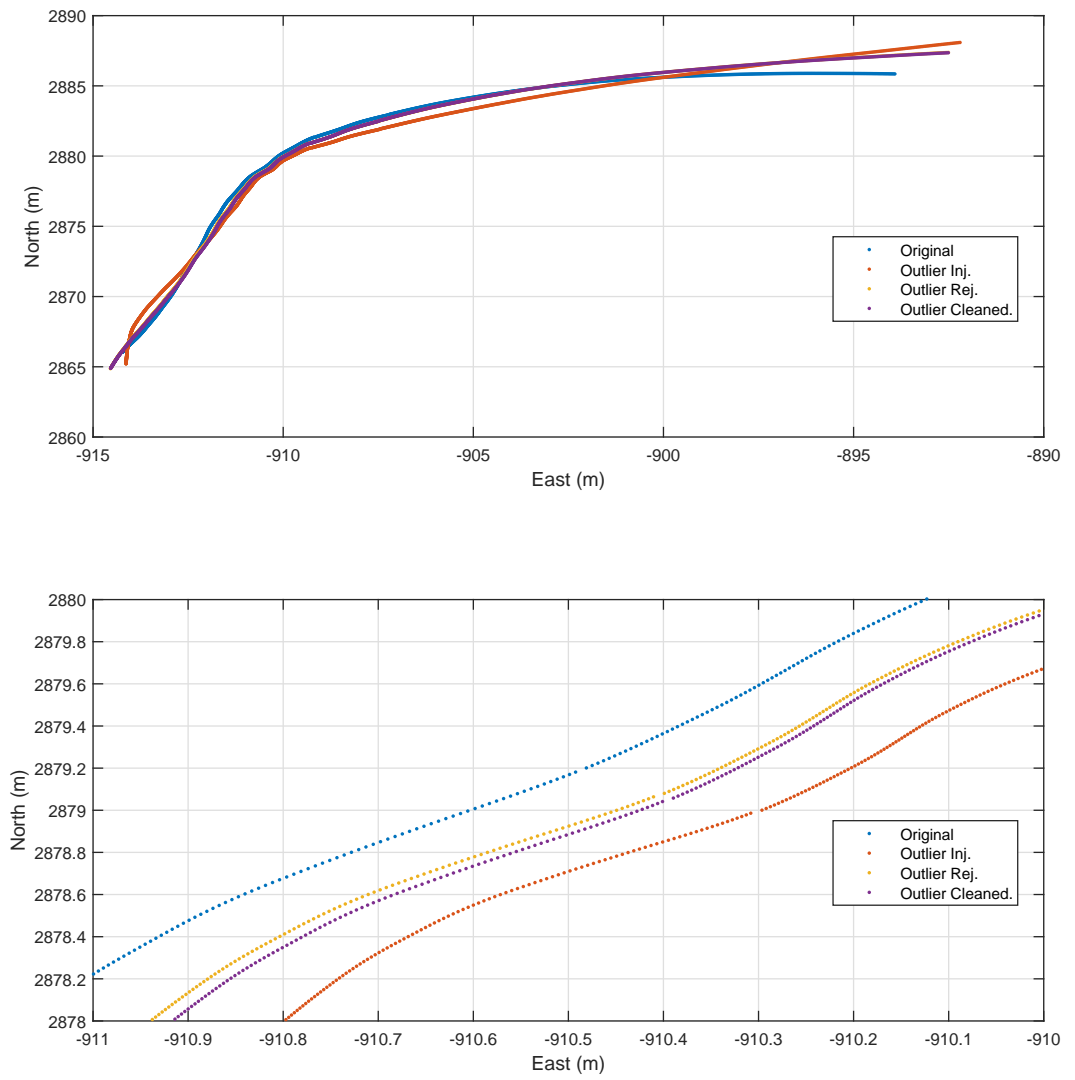
Figure 4.6: Robust GOINS example: GPS-INS positioning results on tangent plane.
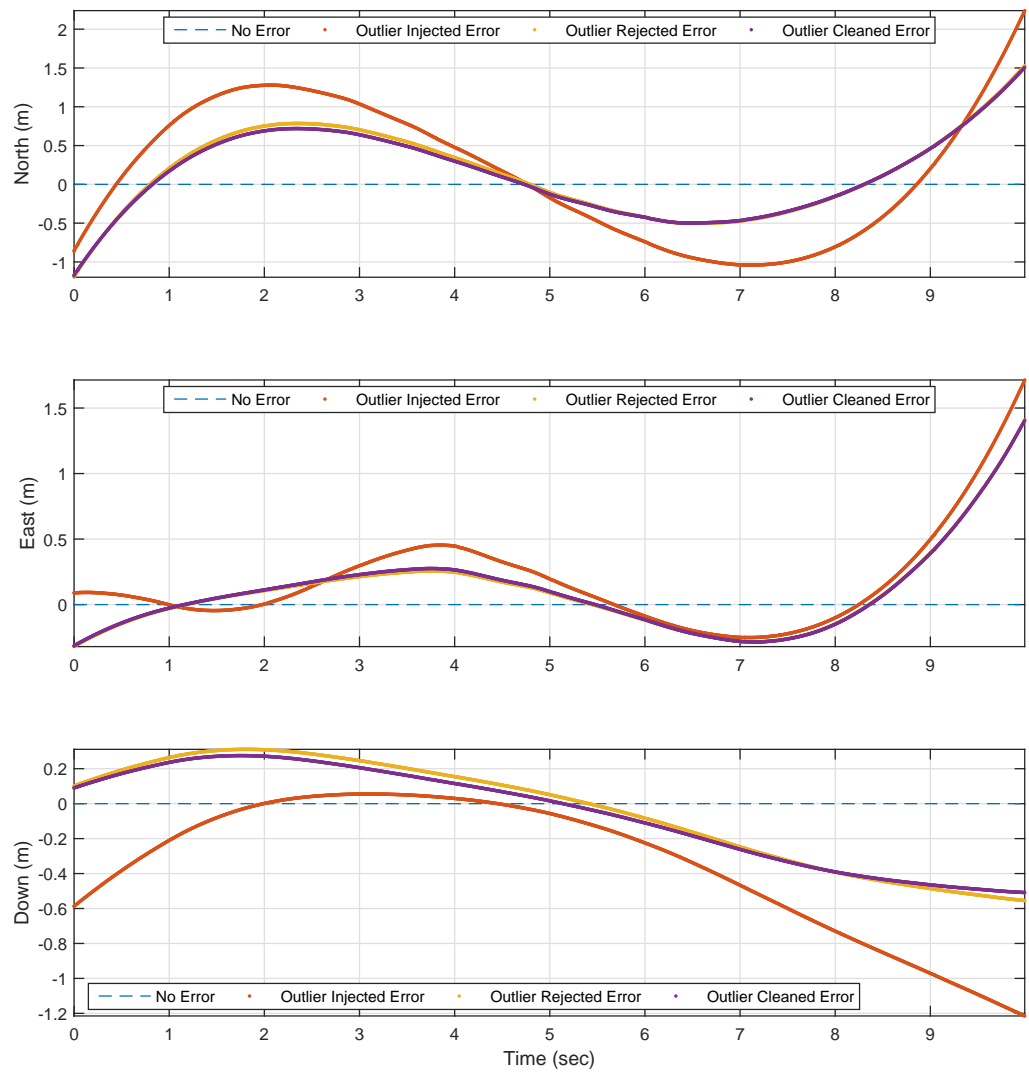
Figure 4.7: Robust GOINS example: GPS-INS 3D positioning results comparison.
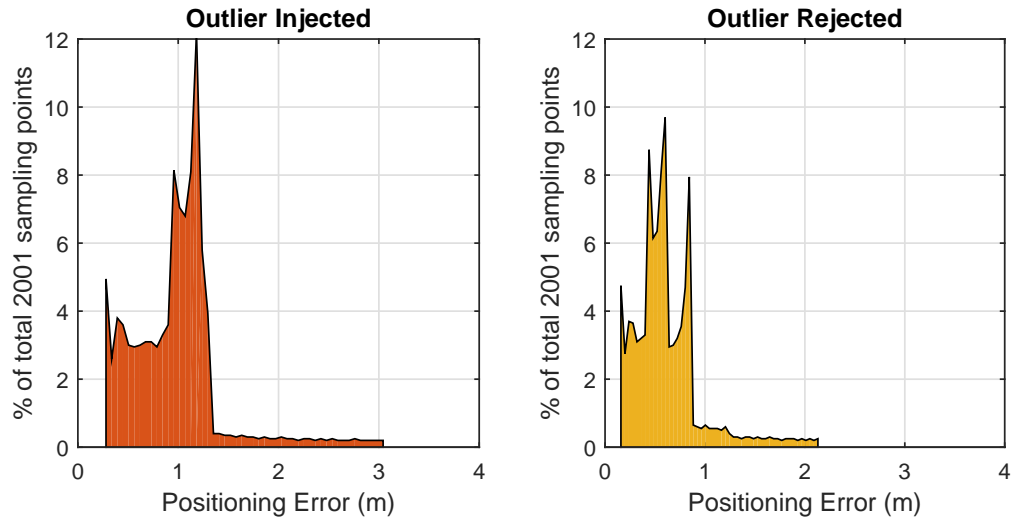
Figure 4.8: Robust GOINS example: GPS-INS 3D positioning error distributions compared with the result from the smoothing with the original data.
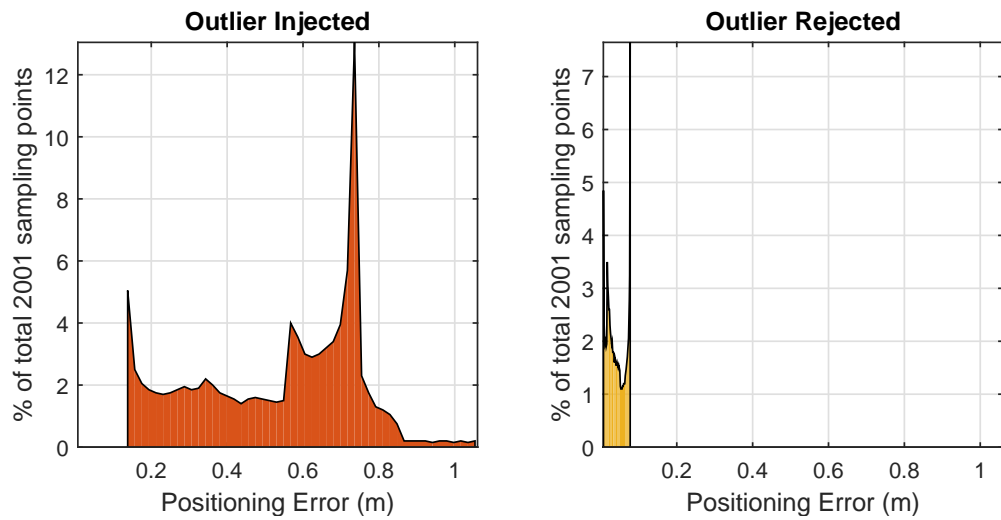
Figure 4.9: Robust GOINS example: GPS-INS 3D positioning error distributions compared with the result from the smoothing with the cleaned data.
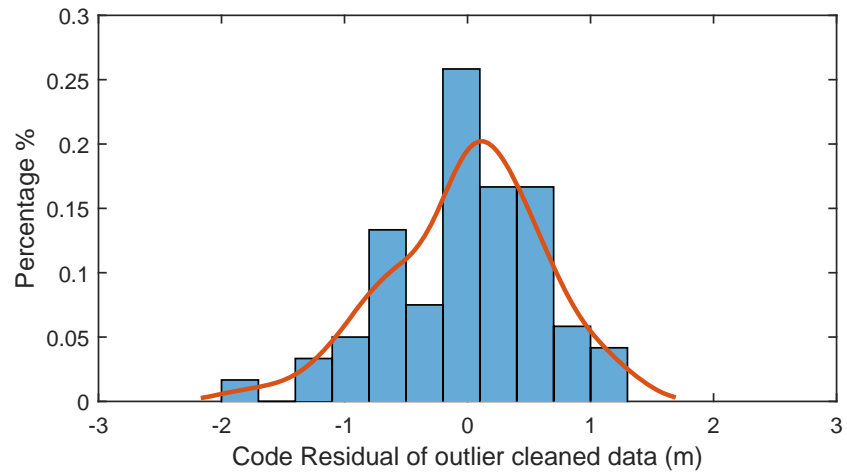
Figure 4.10: Final estimated residuals (weighted) of code measurements in GPS-INS smoothing with the data with outliers cleaned.
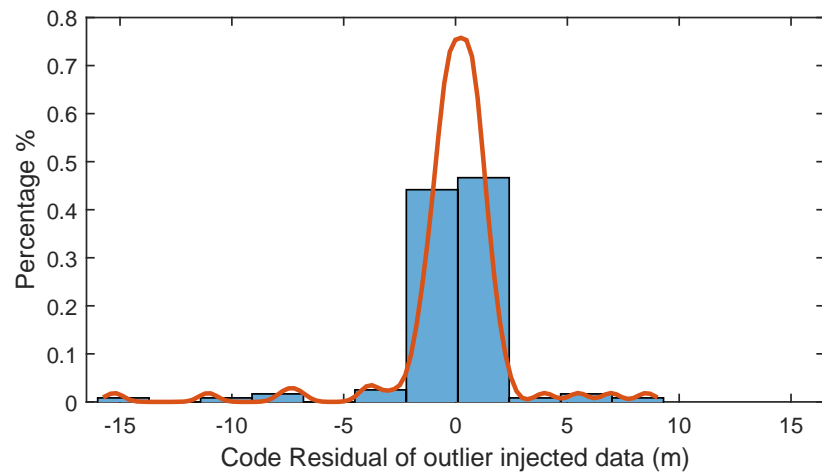
Figure 4.11: Final estimated residuals (weighted) of code measurements in GPS-INS smoothing without any outlier handling.
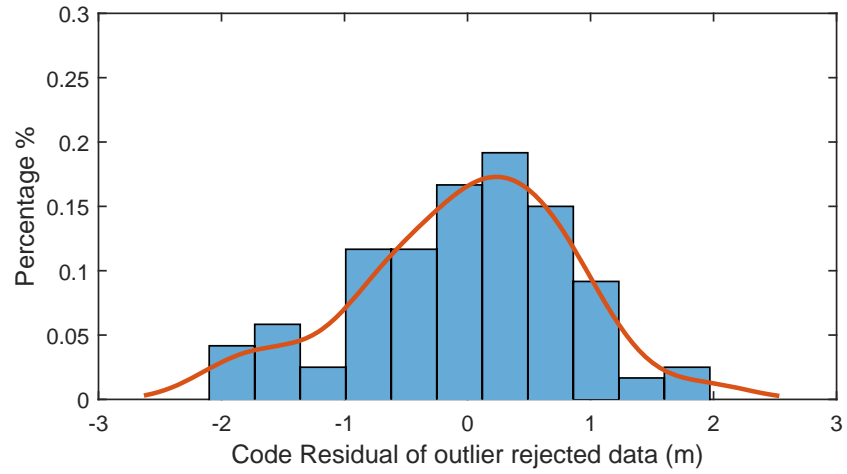
Figure 4.12: Final estimated residuals (weighted) of code measurements in GPS-INS smoothing with HT-LSS outlier rejection.
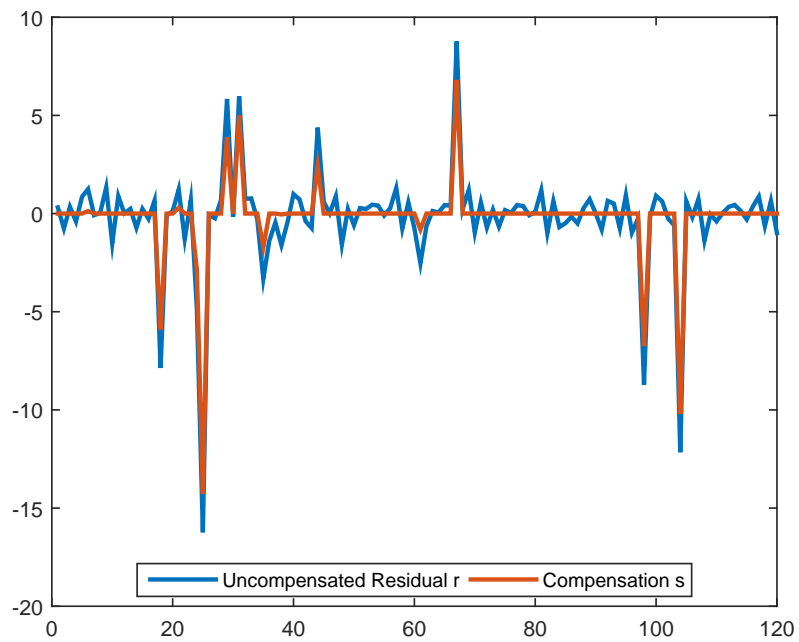
Figure 4.13: Final estimated residuals (weighted) of code measurements in GPS-INS smoothing with HT-LSS outlier rejection.

Chapter 5

Contemplative Real-Time RTK GPS/INS

With the graph optimization based INS proposed in Chapter 3 and the robust optimization technique in Chapter 4, a Contemplative Real-Time (CRT framework for general aided INS and especially for RTK GPS/INS is proposed. Under the CRT framework, conventional EKF and Bayesian smoothing are combined to enhance the accuracy and reliability of the navigation system. Factor Graph model is extended by incorporating integral unknowns to represent the integer ambiguity inherited in the GPS carrier phase measurements for RTK GPS/INS applications. Robust optimization method HT-LSS is applied in the contemplative thread to handling the outliers.

5.1   Integer Ambiguity Resolution

This section extends the discussion of Mixed Integer Least Square (MILS) presented in [11]. In particular, this dissertation proves several propositions giving more insight into the MILS problem. This section first discusses the RTK GPS/INS implementations and the corresponding the Mixed Integer Least Square problem, then the float solution of the integer ambiguity resolution, at last the Integer Least Square is outlined.

5.1.1   RTK GPS/INS

Real-Time Kinematic (RTK) represents the techniques which resolve the integer ambiguity in the GNSS carrier phase measurements, and realize centimeter level positioning accuracy. In this dissertation, RTK in GPS is focused on. Throughout this dissertation, double-differenced GPS measurements are considered [54]. For notational simplicity, it is assumed that the double difference approach completely removes all common-mode errors (e.g., ionosphere, troposphere, satellite clock and ephemeris errors), as well as the receiver clock biases which allows these terms to be dropped throughout the equations. Nonetheless, the remaining errors in these terms will affect the experiment results. However, outlier measurements can exist due to multipath error, heavy foliage, receiver failure, etc.

The (simplified) double-differenced pseudorange (a.k.a. code) measurements for the $i-$th satellite are modeled as

$$\rho^i(t_k) = h_k^i(\boldsymbol{x}(t_k)) + n_\rho^i(t_k), \tag{5.1}$$

where $h_k^i(\boldsymbol{x}(t_k)) = \|\boldsymbol{p}(t_k) - \boldsymbol{p}^i(t_k)\|_2$ is the geometric distance at t_k between the rover position $\boldsymbol{p} \in \mathbb{R}^3$ and the known $i-$th satellite position $\boldsymbol{p}^i \in \mathbb{R}^3$, and $n_\rho^i \sim \mathcal{N}(0, \sigma_\rho^2)$ represents the (non-common mode) measurement noise with standard deviation $\sigma_\rho = 0.1 \sim 3m$, depending on receiver design, environmental factors and the performance of multipath mitigation techniques [12]. In practice, the assumed noise level σ_ρ can vary temporally and spatially for each available satellite. The code measurement in eqn. (5.1) conforms with the standard measurement model in eqn. (3.39).

The (simplified) double-differenced carrier phase measurements for the $i-$th satellite are modeled as

$$\varphi^i(t_k) = h_k^i(\boldsymbol{x}(t_k)) + \lambda N^i(t_k) + n_\varphi^i(t_k), \tag{5.2}$$

where λ is the carrier phase wavelength, and N^i is the unknown integer ambiguity. The measurement noise has distribution $n_\varphi^i \sim \mathcal{N}(0, \sigma_\varphi^2)$. The noise standard deviation σ_φ is millimeter to centimeter level ($< 0.01\sigma_\rho$).

This unknown integer represents the number of whole carrier wave cycles between the satellite and the receiver at the time that phase lock is achieved. If the phase-lock-loop (PLL) in the receiver for the i-th satellite maintains lock without cycle slips on a time interval $[t_1, t_n]$, then the N^i in the carrier phase measurements for

this satellite is a unknown integer constant over this time interval, i.e. $N^i(t_1) = \cdots = N^i(t_n) = N^i$. The receiver reports the lock status to enable detection of such time intervals. The unknown integer must be estimated exactly to enable use of the carrier phase measurement for precise position estimation. Note that the carrier phase measurement model does not match the standard measurement model in eqn. (3.39), because there is a unknown integer variable N^i.

To utilize the carrier phase measurements under conventional EKF framework, the integer ambiguity $\mathbf{N} \triangleq [N^1, \dots, N^m]$ is resolved by the Real-Time Kinematic technique and then the carrier phase measurements can be incorporated into the estimation just like the pseudorange measurements. In particular, the following optimization problem at time step $t = t_k$ is solved.

$$\min_{\boldsymbol{p}(t_k) \in \mathbb{R}^3, \mathbf{N} \in \mathbb{Z}^m} \sum_i \left[\|h_k^i(\boldsymbol{p}(t_k)) - \rho^i(t_k)\|_{\sigma_\rho^2}^2 + \|h_k^i(\boldsymbol{p}(t_k)) + \lambda N^i - \varphi^i(t_k)\|_{\sigma_\varphi^2}^2 \right]. \quad (5.3)$$

Note that in eqn. (5.3) only L1 measurements are considered, but for dual frequency cases wide-lane carrier phase measurements can be formed to facilitate the integer ambiguity resolution from the L1 and L2 carrier phase measurements. In the rest of this section, methods to solve (5.3) is presented.

5.1.2 Mixed Integer Least Square

In terms of the integer ambiguity resolution problem in GNSS research (e.g., Real-Time Kinematic technique), the following Mixed Integer Least Square (MILS)

formulation is often considered to provide optimal solution

$$(\mathbf{X}^*, \mathbf{N}^*) = \underset{\mathbf{X} \in \mathbb{R}^k, \mathbf{N} \in \mathbb{Z}^n}{arg\,min} \|\boldsymbol{r} + \boldsymbol{A}\mathbf{X} + \boldsymbol{B}\mathbf{N}\|^2, \tag{5.4}$$

where $\boldsymbol{r} \in \mathbb{R}^m$ is the measurement residual vector, $\boldsymbol{A}_{m \times k}$ is the Jacobian matrix for the real unknown variable vector $\mathbf{X}$ and $\boldsymbol{B}_{m \times n}$ is the Jacobian matrix for the integral unknown variable vector $\mathbf{N}$. Furthermore, it is necessary that the matrix $[\boldsymbol{A}, \boldsymbol{B}]_{m \times (k+n)}$ is full column rank.

Due to the integral nature of $\mathbf{N}$, eqn. (5.35) cannot be resolved by the Ordinary Least Square (OLS) method [37] which leads to closed solutions. If there is no real unknowns, the problem in eqn. (5.35) is reduced to Integer Least Square (ILS)

$$\underset{\mathbf{N} \in \mathbb{Z}^n}{\min} \|\boldsymbol{r} + \boldsymbol{B}\mathbf{N}\|^2. \tag{5.5}$$

The ILS problem in eqn. (5.38) can be regarded as the generalization of the special ILS considered in the classical LAMBDA method stemmed from [33, 66, 69, 70],

$$\underset{\mathbf{N} \in \mathbb{Z}^n}{\min} \|\mathbf{N} - \hat{\mathbf{N}}\|_{\mathbf{C}_{\hat{\mathbf{N}}}}^2. \tag{5.6}$$

In the method proposed in [11], by letting

$$\boldsymbol{A} = [\boldsymbol{Q}_A, \bar{\boldsymbol{Q}}_A] \begin{bmatrix} \boldsymbol{R}_A \\ \boldsymbol{0} \end{bmatrix}, \tag{5.7}$$

is the QR-decomposition of the matrix $\boldsymbol{A}$, the cost of MILS in eqn. (5.35) is first divided into two components,

$$
\begin{aligned}
& \|\boldsymbol{r} + \boldsymbol{A}\mathbf{X} + \boldsymbol{B}\mathbf{N}\|^2 \\
&= \left\| \begin{bmatrix} \boldsymbol{Q}_A^{\mathsf{T}} \\ \bar{\boldsymbol{Q}}_A^{\mathsf{T}} \end{bmatrix} (\boldsymbol{r} + \boldsymbol{A}\mathbf{X} + \boldsymbol{B}\mathbf{N}) \right\|^2 \\
&= \left\| \begin{bmatrix} \boldsymbol{Q}_A^{\mathsf{T}} \\ \bar{\boldsymbol{Q}}_A^{\mathsf{T}} \end{bmatrix} \boldsymbol{r} + \begin{bmatrix} \boldsymbol{R}_A \\ \mathbf{0} \end{bmatrix} \mathbf{X} + \begin{bmatrix} \boldsymbol{Q}_A^{\mathsf{T}}\boldsymbol{B} \\ \bar{\boldsymbol{Q}}_A^{\mathsf{T}}\boldsymbol{B} \end{bmatrix} \mathbf{N} \right\|^2 \\
&= \|\boldsymbol{Q}_A^{\mathsf{T}}\boldsymbol{r} + \boldsymbol{R}_A\mathbf{X} + \boldsymbol{Q}_A^{\mathsf{T}}\boldsymbol{B}\mathbf{N}\|^2 + \|\bar{\boldsymbol{Q}}_A^{\mathsf{T}}\boldsymbol{r} + \bar{\boldsymbol{Q}}_A^{\mathsf{T}}\boldsymbol{B}\mathbf{N}\|^2.
\end{aligned}
\tag{5.8}
$$

For eqn. (5.37), a crucial observation is that when any integer estimate $\hat{\mathbf{N}}$ is given, there is always a real vector $\hat{\mathbf{X}}$ can be found such that $\|\boldsymbol{Q}_A^{\mathsf{T}}\boldsymbol{r} + \boldsymbol{R}_A\hat{\mathbf{X}} + \boldsymbol{Q}_A^{\mathsf{T}}\boldsymbol{B}\hat{\mathbf{N}}\|^2 = 0$. Thus, the MILS problem in eqn. (5.35) is equivalent to

$$
\mathbf{N}^* = \underset{\mathbf{N}\in\mathbb{Z}^n}{arg\,min} \|\bar{\boldsymbol{Q}}_A^{\mathsf{T}}\boldsymbol{r} + \bar{\boldsymbol{Q}}_A^{\mathsf{T}}\boldsymbol{B}\mathbf{N}\|^2,
\tag{5.9}
$$

and then $\mathbf{X}^*$ can be derived through the back-substitution solving

$$
\boldsymbol{R}_A\mathbf{X} = -\left(\boldsymbol{Q}_A^{\mathsf{T}}\boldsymbol{r} + \boldsymbol{Q}_A^{\mathsf{T}}\boldsymbol{B}\mathbf{N}\right).
$$

By letting $\bar{\boldsymbol{r}} \triangleq \boldsymbol{Q}_A^{\mathsf{T}}\boldsymbol{r}$ and $\bar{\boldsymbol{B}} \triangleq \bar{\boldsymbol{Q}}_A^{\mathsf{T}}\boldsymbol{B}$, eqn. (5.9) can be rewritten as

$$
\mathbf{N}^* = \underset{\mathbf{N}\in\mathbb{Z}^n}{arg\,min} \|\bar{\boldsymbol{r}} + \bar{\boldsymbol{B}}\mathbf{N}\|^2
\tag{5.10}
$$

which matches the standard form of ILS in eqn. (5.38). The solution of the ILS is introduced in Section 5.1.4.

5.1.3 Float solution

Alternatively to direct solving (5.10) by ILS, the **float solution** is often derived as an intermediate step for integer ambiguity resolution,

$$\check{\mathbf{N}} = \underset{\mathbf{N} \in \mathbb{R}^n}{arg\,min} \|\bar{\boldsymbol{r}} + \bar{\boldsymbol{B}}\mathbf{N}\|^2, \tag{5.11}$$

where the integral nature of $\mathbf{N}$ is neglected and treated as a real vector. Thus, the float solution can be derived by the closed form of Ordinary Least Square (OLS). In particular, letting

$$\boldsymbol{B} = [\boldsymbol{Q}_{\bar{B}}, \bar{\boldsymbol{Q}}_{\bar{B}}] \begin{bmatrix} \boldsymbol{R}_{\bar{B}} \\ \boldsymbol{0} \end{bmatrix}, \tag{5.12}$$

then the float solution is

$$\check{\mathbf{N}} = -\left(\bar{\boldsymbol{B}}^{\mathsf{T}}\bar{\boldsymbol{B}}\right)^{-1}\bar{\boldsymbol{B}}^{\mathsf{T}}\bar{\boldsymbol{r}} = -\boldsymbol{R}_{\bar{B}}^{-1}\boldsymbol{Q}_{\bar{B}}^{\mathsf{T}}\bar{\boldsymbol{r}}, \tag{5.13}$$

and the corresponding covariance is

$$\mathbf{C}_{\check{\mathbf{N}}} = \left(\bar{\boldsymbol{B}}^{\mathsf{T}}\bar{\boldsymbol{B}}\right)^{-1} = \left(\boldsymbol{B}^{\mathsf{T}}\bar{\boldsymbol{Q}}_A\bar{\boldsymbol{Q}}_A^{\mathsf{T}}\boldsymbol{B}\right)^{-1}. \tag{5.14}$$

This float solution is often used in integer validation, e.g. in the ratio test [68]. Also, based on the following proposition, the original integer ambiguity resolution in eqn. (5.35) can be reformed as ILS problem when the float solution is given.

Proposition 5.1. *The optimal integer estimate* $\mathbf{N}^*$ *in eqn. (5.35) can also be resolved through*

$$\mathbf{N}^* = \underset{\mathbf{N} \in \mathbb{Z}^n}{arg\,min} \|\mathbf{N} - \check{\mathbf{N}}\|^2_{\mathbf{C}_{\check{\mathbf{N}}}}, \tag{5.15}$$

107

where $\check{\mathbf{N}}$ and $\mathbf{C}_{\check{\mathbf{N}}}$ are the float solution and the corresponding covariance matrix derived in eqn. (5.13) and (5.14).

Proof. Based on the analysis in Section 5.1.2, it is sufficient to show that

$$\underset{\mathbf{N}\in\mathbb{Z}^n}{arg\,min}\,\|\mathbf{N} - \check{\mathbf{N}}\|^2_{\mathbf{C}_{\check{\mathbf{N}}}} = \underset{\mathbf{N}\in\mathbb{Z}^n}{arg\,min}\,\|\bar{r} + \bar{\boldsymbol{B}}\mathbf{N}\|^2. \qquad (5.16)$$

With eqn. (5.12), (5.13) and (5.14), it follows that

$$\left\|\mathbf{N} - \check{\mathbf{N}}\right\|^2_{\mathbf{C}_{\check{\mathbf{N}}}}$$

$$= \left(\mathbf{N} - \check{\mathbf{N}}\right)^\intercal \left(\bar{\boldsymbol{B}}^\intercal \bar{\boldsymbol{B}}\right) \left(\mathbf{N} - \check{\mathbf{N}}\right)$$

$$= \left\|\bar{\boldsymbol{B}}\mathbf{N} - \bar{\boldsymbol{B}}\check{\mathbf{N}}\right\|^2$$

$$= \left\|\bar{\boldsymbol{B}}\mathbf{N} + \bar{\boldsymbol{B}}\boldsymbol{R}_{\bar{B}}^{-1}\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r}\right\|^2$$

$$= \left\|\bar{\boldsymbol{B}}\mathbf{N} + \left[\boldsymbol{Q}_{\bar{B}}, \bar{\boldsymbol{Q}}_{\bar{B}}\right] \begin{bmatrix} \boldsymbol{R}_{\bar{B}} \\ \mathbf{0} \end{bmatrix} \boldsymbol{R}_{\bar{B}}^{-1}\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r}\right\|^2$$

$$= \left\|\bar{\boldsymbol{B}}\mathbf{N} + \boldsymbol{Q}_{\bar{B}}\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r}\right\|^2$$

$$= \left(\bar{\boldsymbol{B}}\mathbf{N} + \boldsymbol{Q}_{\bar{B}}\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r}\right)^\intercal \left(\bar{\boldsymbol{B}}\mathbf{N} + \boldsymbol{Q}_{\bar{B}}\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r}\right)$$

$$= \|\bar{\boldsymbol{B}}\mathbf{N}\|^2 + 2\mathbf{N}^\intercal\bar{\boldsymbol{B}}^\intercal\boldsymbol{Q}_{\bar{B}}\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r} + \|\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r}\|^2$$

$$= \|\bar{\boldsymbol{B}}\mathbf{N}\|^2 + 2\mathbf{N}^\intercal \left[\boldsymbol{R}_{\bar{B}}^\intercal, \mathbf{0}\right] \begin{bmatrix} \boldsymbol{Q}_{\bar{B}}^\intercal \\ \bar{\boldsymbol{Q}}_{\bar{B}}^\intercal \end{bmatrix} \boldsymbol{Q}_{\bar{B}}\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r} + \|\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r}\|^2$$

$$= \|\bar{\boldsymbol{B}}\mathbf{N}\|^2 + 2\mathbf{N}^\intercal\boldsymbol{R}_{\bar{B}}^\intercal\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r} + \|\boldsymbol{Q}_{\bar{B}}^\intercal\bar{r}\|^2, \qquad (5.17)$$

where the last equity is from $\boldsymbol{Q}_{\bar{B}}^\intercal\boldsymbol{Q}_{\bar{B}} = \boldsymbol{I}$.

On the other hand, we have that

$$\left\|\bar{\boldsymbol{r}} + \bar{\boldsymbol{B}}\mathbf{N}\right\|^2$$

$$= \left\|\begin{bmatrix} \boldsymbol{Q}_{\bar{B}}^\mathsf{T} \\ \bar{\boldsymbol{Q}}_{\bar{B}}^\mathsf{T} \end{bmatrix} (\bar{\boldsymbol{r}} + \bar{\boldsymbol{B}}\mathbf{N})\right\|^2$$

$$= \left\|\boldsymbol{Q}_{\bar{B}}^\mathsf{T}\bar{\boldsymbol{r}} + \boldsymbol{R}_{\bar{B}}\mathbf{N}\right\|^2 + \left\|\bar{\boldsymbol{Q}}_{\bar{B}}^\mathsf{T}\bar{\boldsymbol{r}}\right\|^2.$$

Since $\left\|\bar{\boldsymbol{Q}}_{\bar{B}}^\mathsf{T}\bar{\boldsymbol{r}}\right\|^2$ is a constant, it is equivalent to minimize $\left\|\boldsymbol{Q}_{\bar{B}}^\mathsf{T}\bar{\boldsymbol{r}} + \boldsymbol{R}_{\bar{B}}\mathbf{N}\right\|^2$. Then, it follows that

$$\left\|\boldsymbol{Q}_{\bar{B}}^\mathsf{T}\bar{\boldsymbol{r}} + \boldsymbol{R}_{\bar{B}}\mathbf{N}\right\|^2$$

$$= \mathbf{N}^\mathsf{T}\boldsymbol{R}_{\bar{B}}^\mathsf{T}\boldsymbol{R}_{\bar{B}}\mathbf{N} + 2\mathbf{N}^\mathsf{T}\boldsymbol{R}_{\bar{B}}^\mathsf{T}\boldsymbol{Q}_{\bar{B}}^\mathsf{T}\bar{\boldsymbol{r}} + \left\|\boldsymbol{Q}_{\bar{B}}^\mathsf{T}\bar{\boldsymbol{r}}\right\|^2$$

$$= \left\|\bar{\boldsymbol{B}}\mathbf{N}\right\|^2 + 2\mathbf{N}^\mathsf{T}\boldsymbol{R}_{\bar{B}}^\mathsf{T}\boldsymbol{Q}_{\bar{B}}^\mathsf{T}\bar{\boldsymbol{r}} + \left\|\boldsymbol{Q}_{\bar{B}}^\mathsf{T}\bar{\boldsymbol{r}}\right\|^2,$$

where the last equity is from $\boldsymbol{R}_{\bar{B}}^\mathsf{T}\boldsymbol{R}_{\bar{B}} = \bar{\boldsymbol{B}}^\mathsf{T}\bar{\boldsymbol{B}}$. Thus, from eqn. (5.17), it follows that minimizing $\|\mathbf{N} - \check{\mathbf{N}}\|_{\mathbf{C}_{\check{\mathbf{N}}}}^2$ is equivalent to minimizing $\|\bar{\boldsymbol{r}} + \bar{\boldsymbol{B}}\mathbf{N}\|^2$, and this concludes the proof. $\qquad\square$

Besides of eqn. (5.11), another float solution can be resolved from the following **information form** method. Later, Proposition 5.4 shows that this float solution from the information formulation is equivalent to that from The original cost in eqn.

(5.35) can be expanded as

$$\|\boldsymbol{r} + \boldsymbol{A}\mathbf{X} + \boldsymbol{B}\mathbf{N}\|^2$$
$$= \|\boldsymbol{r}\|^2 + 2\boldsymbol{r}^\mathsf{T}\,[\boldsymbol{A}, \boldsymbol{B}]\begin{bmatrix}\boldsymbol{X}\\\boldsymbol{N}\end{bmatrix} + [\boldsymbol{X}^\mathsf{T}, \boldsymbol{N}^\mathsf{T}]\begin{bmatrix}\boldsymbol{A}^\mathsf{T}\boldsymbol{A} & \boldsymbol{A}^\mathsf{T}\boldsymbol{B}\\\boldsymbol{B}^\mathsf{T}\boldsymbol{A} & \boldsymbol{B}^\mathsf{T}\boldsymbol{B}\end{bmatrix}\begin{bmatrix}\boldsymbol{X}\\\boldsymbol{N}\end{bmatrix}. \qquad (5.18)$$

By treating $\boldsymbol{N}$ as an real vector, the optimality condition leads to solving the linear system

$$\underbrace{\begin{bmatrix}\boldsymbol{A}^\mathsf{T}\boldsymbol{A} & \boldsymbol{A}^\mathsf{T}\boldsymbol{B}\\\boldsymbol{B}^\mathsf{T}\boldsymbol{A} & \boldsymbol{B}^\mathsf{T}\boldsymbol{B}\end{bmatrix}}_{\boldsymbol{\Lambda}}\begin{bmatrix}\boldsymbol{X}\\\boldsymbol{N}\end{bmatrix} = -\begin{bmatrix}\boldsymbol{A}^\mathsf{T}\boldsymbol{r}\\\boldsymbol{B}^\mathsf{T}\boldsymbol{r}\end{bmatrix}. \qquad (5.19)$$

Note that in eqn. (5.19), $\boldsymbol{\Lambda}$ is the corresponding information matrix. Schur complement [23, 25, 41] can be used to solve the above system by marginalizing the real unknowns $\boldsymbol{X}$,

$$\left(\boldsymbol{B}^\mathsf{T}\boldsymbol{B} - \boldsymbol{B}^\mathsf{T}\boldsymbol{A}\left(\boldsymbol{A}^\mathsf{T}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^\mathsf{T}\boldsymbol{B}\right)\boldsymbol{N} = -\boldsymbol{B}^\mathsf{T}\boldsymbol{r} + \boldsymbol{B}^\mathsf{T}\boldsymbol{A}\left(\boldsymbol{A}^\mathsf{T}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^\mathsf{T}\boldsymbol{r}. \qquad (5.20)$$

Letting

$$\boldsymbol{Q}_p \triangleq \left(\boldsymbol{I} - \boldsymbol{A}\left(\boldsymbol{A}^\mathsf{T}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^\mathsf{T}\right) \in \mathbb{R}^{m \times m}, \qquad (5.21)$$

then eqn. (5.20) can be rewritten as

$$\boldsymbol{B}^\mathsf{T}\boldsymbol{Q}_p\boldsymbol{B}\boldsymbol{N} = -\boldsymbol{B}^\mathsf{T}\boldsymbol{Q}_p\boldsymbol{r}. \qquad (5.22)$$

Note that $\boldsymbol{Q}_p$ is a projection matrix onto the subspace of $\mathbb{R}^m$ which is the orthogonal complement of the range space $\mathcal{C}(\boldsymbol{A})$ of $\boldsymbol{A}$ [37], i.e., the left nullspace of $\boldsymbol{A}$ which denoted as $\mathcal{LN}(\boldsymbol{A})$ or $\mathcal{N}(\boldsymbol{A}^\mathsf{T})$. Furthermore, along with that $[\boldsymbol{A}, \boldsymbol{B}]$ has full column

rank, it can be proved that $\boldsymbol{B}^\mathsf{T}\boldsymbol{Q}_p\boldsymbol{B}$ has full rank and then invertible. A useful *lemma* can be proved to facilitate the later proof.

Lemma 5.2. *With the definition of eqn. (5.21) and (5.7), it follows that*

$$\boldsymbol{Q}_p = \bar{\boldsymbol{Q}}_A\bar{\boldsymbol{Q}}_A^\mathsf{T}.$$

Proof. With the definition (5.21) and (5.7),it follows that

$$
\begin{aligned}
\boldsymbol{Q}_p &= \boldsymbol{I} - [\boldsymbol{Q}_A, \bar{\boldsymbol{Q}}_A] \begin{bmatrix} \boldsymbol{R}_A \\ \boldsymbol{0} \end{bmatrix} \left([\boldsymbol{R}_A^\mathsf{T}, \boldsymbol{0}^\mathsf{T}] \begin{bmatrix} \boldsymbol{Q}_A^\mathsf{T} \\ \bar{\boldsymbol{Q}}_A^\mathsf{T} \end{bmatrix} [\boldsymbol{Q}_A, \bar{\boldsymbol{Q}}_A] \begin{bmatrix} \boldsymbol{R}_A \\ \boldsymbol{0} \end{bmatrix} \right)^{-1} [\boldsymbol{R}_A^\mathsf{T}, \boldsymbol{0}^\mathsf{T}] \begin{bmatrix} \boldsymbol{Q}_A^\mathsf{T} \\ \bar{\boldsymbol{Q}}_A^\mathsf{T} \end{bmatrix} \\
&= \boldsymbol{I} - \boldsymbol{Q}_A\boldsymbol{R}_A \left(\boldsymbol{R}_A^\mathsf{T}\boldsymbol{R}_A \right)^{-1} \boldsymbol{R}_A^\mathsf{T}\boldsymbol{Q}_A^\mathsf{T} \\
&= \boldsymbol{I} - \boldsymbol{Q}_A\boldsymbol{R}_A\boldsymbol{R}_A^{-1} \left(\boldsymbol{R}_A^\mathsf{T} \right)^{-1} \boldsymbol{R}_A^\mathsf{T}\boldsymbol{Q}_A^\mathsf{T} \\
&= [\boldsymbol{Q}_A, \bar{\boldsymbol{Q}}_A] \begin{bmatrix} \boldsymbol{Q}_A^\mathsf{T} \\ \bar{\boldsymbol{Q}}_A^\mathsf{T} \end{bmatrix} - \boldsymbol{Q}_A\boldsymbol{Q}_A^\mathsf{T} \\
&= \bar{\boldsymbol{Q}}_A\bar{\boldsymbol{Q}}_A^\mathsf{T}.
\end{aligned}
$$

$\square$

Remark 5.3. *Intuitively, $\boldsymbol{I} - \boldsymbol{A}\left(\boldsymbol{A}^\mathsf{T}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^\mathsf{T}$ is a projection onto the left nullspace of $\boldsymbol{A}$, i.e., $\mathcal{N}(\boldsymbol{A}^\mathsf{T})$, while the columns of $\bar{\boldsymbol{Q}}_A$ span the left nullspace of $\boldsymbol{A}$. Thus, $\bar{\boldsymbol{Q}}_A\bar{\boldsymbol{Q}}_A^\mathsf{T}$ is also a projection onto $\mathcal{N}(\boldsymbol{A}^\mathsf{T})$.* $\triangle$

Then, we have the following proposition to conclude the discussion in this subsection.

Proposition 5.4. *The integer estimates derived from the following optimizations are equal to each other,*

1. *MILS in eqn. (5.35)*

$$\min_{\mathbf{X} \in \mathbb{R}^k, \mathbf{N} \in \mathbb{Z}^n} \|\boldsymbol{r} + \boldsymbol{A}\mathbf{X} + \boldsymbol{B}\mathbf{N}\|^2,$$

2. *ILS in eqn. (5.38)*

$$\min_{\mathbf{N} \in \mathbb{Z}^n} \|\bar{\boldsymbol{r}} + \bar{\boldsymbol{B}}\mathbf{N}\|^2,$$

3. *ILS with the float solution in eqn. (5.15)*

$$\min_{\mathbf{N} \in \mathbb{Z}^n} \|\mathbf{N} - \check{\mathbf{N}}\|^2_{\mathbf{C}_{\check{\mathbf{N}}}},$$

4. *ILS from the information form (5.22)*

$$\min_{\mathbf{N} \in \mathbb{Z}^n} \|\boldsymbol{B}^\mathsf{T}\boldsymbol{Q}_p\boldsymbol{B}\boldsymbol{N} + \boldsymbol{B}^\mathsf{T}\boldsymbol{Q}_p\boldsymbol{r}\|^2.$$

Proof. The equivalence of item *1,2,3* is well discussed in Proposition 5.1, so it is sufficient to prove item *3* and item *4* are equivalent. With eqn. (5.22) and that $\boldsymbol{B}^\mathsf{T}\boldsymbol{Q}_p\boldsymbol{B}$ is invertible, another float solution can be derived from the information form

$$\check{\mathbf{N}} = -\left(\boldsymbol{B}^\mathsf{T}\boldsymbol{Q}_p\boldsymbol{B}\right)^{-1}\boldsymbol{B}^\mathsf{T}\boldsymbol{Q}_p\bar{\boldsymbol{r}}, \tag{5.23}$$

with the corresponding covariance matrix

$$\mathbf{C}_{\check{\mathbf{N}}} = -\left(\boldsymbol{B}^\mathsf{T}\boldsymbol{Q}_p\boldsymbol{B}\right)^{-1}. \tag{5.24}$$

Then, along with Lemma 5.2, it follows that the two float solutions are equal

$$\underset{\mathbf{N}\in\mathbb{Z}^n}{arg\,min}\|\mathbf{N}-\check{\mathbf{N}}\|_{\mathbf{C}_{\check{\mathbf{N}}}}^2 = \underset{\mathbf{N}\in\mathbb{Z}^n}{arg\,min}\|\mathbf{N}-\check{\mathbf{N}}\|_{\mathbf{C}_{\check{\mathbf{N}}}}^2.$$

Similar to eqn. (5.16), it can be proven that

$$\underset{\mathbf{N}\in\mathbb{Z}^n}{arg\,min}\|\mathbf{N}-\check{\mathbf{N}}\|_{\mathbf{C}_{\check{\mathbf{N}}}}^2 = \underset{\mathbf{N}\in\mathbb{Z}^n}{arg\,min}\|\boldsymbol{B}^{\mathsf{T}}\boldsymbol{Q}_p\boldsymbol{B}\boldsymbol{N} + \boldsymbol{B}^{\mathsf{T}}\boldsymbol{Q}_p\boldsymbol{r}\|^2,$$

and this concludes this proof. □

5.1.4 Integer Least Square

Integer Least Square (ILS) considers a problem in the following form,

$$\min_{\boldsymbol{z}\in\mathbb{Z}^n}\|\boldsymbol{y}-\boldsymbol{B}\boldsymbol{z}\|^2, \tag{5.25}$$

where $\boldsymbol{y}\in\mathbb{R}^m$. The state-of-the-art Integer Least Square (ILS) solution can be divided into two steps: reduction and search. The reduction step is also referred as decorrelation, whose aim is to facilitate the later search step. Since this dissertation does not make contribution to ILS, the basic idea of reduction and search is briefly reviewed in this section. More details can be found from reference [2, 9, 11, 66, 69, 70].

The reduction process which transform $\boldsymbol{B}$ matrix into a upper-triangular one can be done through LLL reduction [42, 83]. Furthermore, the LLL reduction can be referred as a QRZ factorization:

$$\boldsymbol{Q}\boldsymbol{B}\boldsymbol{Z} = \begin{bmatrix}\boldsymbol{R}\\\boldsymbol{0}\end{bmatrix}, \quad \text{or } \boldsymbol{B} = \boldsymbol{Q}\begin{bmatrix}\boldsymbol{R}\\\boldsymbol{0}\end{bmatrix}\boldsymbol{Z}^{-1}, \tag{5.26}$$

where $\boldsymbol{Q} = \begin{bmatrix} \underbrace{\boldsymbol{Q}_1}_{n}, \underbrace{\boldsymbol{Q}_2}_{m-n} \end{bmatrix} \in \mathbb{R}^{m \times m}$ is orthogonal, $\boldsymbol{Z}$ is unimodual[1], and $\boldsymbol{R} \in \mathbb{R}^{n \times n}$ is nonsingular upper-triangular and satisfies the following two LLL reduction criteria [42]:

$$|r_{ij}| \leq \frac{1}{2}|r_{ii}|, \quad r_{ii}^2 \leq r_{i,i+1}^2 + r_{i+1,i+1}^2$$

for $i = 1, \dots, n-1, j = i+1, \dots, n$. With the QRZ factorization in eqn. (5.26), the original ILS can be decomposed as

$$\|\boldsymbol{y} - \boldsymbol{B}\boldsymbol{z}\|^2 = \|\boldsymbol{Q}_1^\mathsf{T}\boldsymbol{y} - \boldsymbol{R}\boldsymbol{Z}^{-1}\boldsymbol{z}\|^2 + \|\boldsymbol{Q}_2^\mathsf{T}\boldsymbol{y}\|^2.$$

Then, by defining $\bar{\boldsymbol{y}} \triangleq \boldsymbol{Q}_1^\mathsf{T}\boldsymbol{y}$ and $\bar{\boldsymbol{z}} = \boldsymbol{Z}^{-1}\boldsymbol{z}$, the original ILS (5.25) is transformed into a new one,

$$\min_{\bar{\boldsymbol{z}} \in \mathbb{Z}^n} \|\bar{\boldsymbol{y}} - \boldsymbol{R}\bar{\boldsymbol{z}}\|^2.$$

After the above reduction, search process can use the Schnorr-Euchner enumeration strategy based algorithm from Lattice Theory research [2] to enumerate possible $\bar{\boldsymbol{z}} \in \mathbb{Z}^n$. Let us assume that we want to find the best p optimal solutions to ((5.25)). Then, the search algorithm first find p integer points $\{\bar{\boldsymbol{z}}^{(i)}\}_{i=1}^p$ with increasing costs. Furthermore, the largest cost is defined as

$$\beta \triangleq \|\bar{\boldsymbol{y}} - \boldsymbol{R}\bar{\boldsymbol{z}}^{(p)}\|^2.$$

Then, the algorithm searches the ellipsoid

$$\|\bar{\boldsymbol{y}} - \boldsymbol{R}\bar{\boldsymbol{z}}\|^2 < \beta$$

---

[1] $\boldsymbol{Z}$ is an integer matrix and $|det(\boldsymbol{Z})| = 1$ thus $\boldsymbol{Z}^{-1}$ is also an integer matrix.

to find a new integer point. When a new integer point is found, $\bar{\boldsymbol{z}}^{(p)}$ is removed and the new point is inserted into the sequence $\{\bar{\boldsymbol{z}}^{(i)}\}_{i=1}^{p-1}$, maintaining the increasing order of the corresponding costs. Then, the ellipsoid is shrunk with the new $\bar{\boldsymbol{z}}^{(p)}$ and the search algorithm searches a new integer point within the new ellipsoid. This process repeats until no new integer point can be found.

5.2 Contemplative Real-Time Approach for GPS/INS

This work develops a Contemplative Real-Time (CRT) approach proposed in [14–16, 81, 82] for the GPS/INS application. The CRT approach has both real-time and contemplative aspects. The real-time state estimate is required for control and planning purposes, without latency. The contemplative aspects are intended to enhance accuracy through implementing the robust graph optimization approach for INS presented in Chapter 3 and Chapter 4.

The real-time aspect of the proposed CRT framework can be standard EKF which is often implemented for most GPS/INS applications. For more details of EKF based GPS/INS, see Section 3.1.5. In this chapter only the contemplative aspect of the CRT framework is discussed in details.

The CRT framework considers a trajectory estimation problem over a time interval $[t_1, t_K]$, where t_K is the most recent GPS measurement time. The interval $[t_1, t_K]$ is
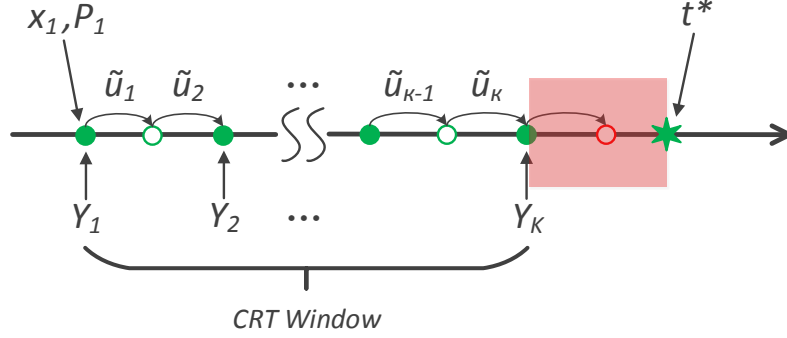
Figure 5.1: Measurement timeline for one CRT cycle.

referred as the *CRT window*. This window contains K GPS measurement epochs, where K can be designer specified, time varying, or data dependent. A typical, but simplified, measurement scenario for $[t_1, t_K]$ is depicted in Fig. 5.1. The dots on the time-line indicate IMU measurement times τ_n. Typically the number of IMU measurements between GPS measurements is very high (i.e. $t_{i+1} - t_i \gg \tau_{j+1} - \tau_j$), for two reasons. First, the IMU sample frequency (e.g. $200Hz$) is at least twice of the IMU bandwidth, which is higher than the vehicle motion bandwidth. Second, GPS sample rate ($\leq 10Hz$) is usually much lower than that of IMU. The state transition between these times is constrained by the kinematic model of eqn. (3.2) and the IMU data. Additional constraints are imposed by the initial estimate $(\boldsymbol{x}_1, \mathbf{P}_1)$ depicted above the initial state, and GPS measurements depicted below the time-line. Each of these constraints is quantified by a probability density that enables a Bayesian estimation formulation for the CRT estimation problem. While

116

Fig. 5.1 depicts all GPS measurements occurring at the IMU measurement time, unaligned measurements can be addressed by interpolation, and unknown latencies can be calibrated by the methods in [45].

The contemplative process starts when $t = t_K$ when the K-th epoch GPS measurements $\mathbf{Y}_K$ are received. At t_K, all IMU, GPS are available for the time interval $[t_1, t_K]$. A prior for the initial state $\boldsymbol{x}(t_1) \sim \mathcal{N}(\boldsymbol{x}_1, \mathbf{P}_1)$ is also available. Starting from the real-time estimates, the CRT algorithm will contemplate the available information to reliably and accurately compute the state trajectory over the CRT window using the robust graph optimization method for INS shown in Chapter 3 and Chapter 4. This contemplative process ends at a time $t^* > t_K$, ideally providing an optimal trajectory estimate $\hat{\boldsymbol{x}}(t)$ for $t \in [t_1, t_K]$ from which the effects of outlier measurements have been removed. For the computation time interval $t \in [t_K, t^*]$, the real-time estimate of the realtime state estimate $\hat{\boldsymbol{x}}(t)$ is maintained by the INS using the IMU data and starting from the prior estimate of $\boldsymbol{x}(t_K)$. At $t = t^*$, $\hat{\boldsymbol{x}}(t_K)$ is corrected to the result of the CRT contemplative process and propagated through time using the IMU data and eqn. (3.2) to provide an improved estimate of $\boldsymbol{x}(t)$ at the present time. At some time $t \geq t^*$, the CRT window can be redefined and the process can repeat indefinitely.

The proposed CRT framework do not only apply to general aided INS implementations (e.g. DGPS/INS), but also applies to RTK GPS/INS. However, for a RTK GPS/INS navigation system, the smoothing estimation problem is a bit different from

117

the general one defined in Chapter 3, due to existence of the integer ambiguity. In Section 5.3, the classical factor graph modeling is extended by incorporate integral unknowns, and an efficient solution to corresponding Nonlinear Mixed Integer Least Square problem is proposed.

5.3 Probabilistic Graphical Modeling for RTK G-PS/INS

In this work, the estimation problem of RTK GPS/INS is formalized under the Robust Graph Optimization INS (GOINS) framework presented in Chapter 3 and Chapter 4. For the simplicity of discussion in this paper, the following assumption is made.

Assumption 5.5. *For $t \in [t_1, t_K]$, the receiver provides valid carrier phase measurements for m satellites, without loss of lock. Then, the unknown integers in the carrier phase measurements from these m satellites are constants over $[t_1, t_K]$.* $\triangle$

Remark 5.6. *Assumption 6.1 is made for the convenience of presenting the main ideas in this article. The proposed framework can be easily extended to more complicated application scenarios. In fact, it is straightforward to either ensure the assumption is true, by decreasing m or K. Decreasing either parameter affects both the achievable accuracy and reliability. Also, GPS measurements occur at 1 Hz, which is higher than*

necessary to maintain the desired decimeter accuracy, so it is also straightforward to
simply not perform the algorithm on those intervals when the assumption is not valid
for a sufficiently large m. △

Under Assumption 6.1, the RTK GPS/INS estimation problem considered in this chapter can be stated as follows.

For a system described by eqn. (3.1), we have

- an initial distribution for the state $\boldsymbol{x}(t_1) \sim \mathcal{N}(\boldsymbol{x}_1, \mathbf{P}_1)$,

- IMU measurements $\mathbf{U} = \{\mathbf{U}_k\}_{k=1}^{K-1}$, where

$$\mathbf{U}_k = \{\tilde{\boldsymbol{u}}(\tau_n), t_k \leq \tau_n \leq t_{k+1}\},$$

- DGPS code and carrier phase measurements $\mathbf{Y} = \{\mathbf{Y}_k\}_{k=1}^{K}$, where

$$\mathbf{Y}_k = \{\boldsymbol{\rho}^i(t_k)\}_{i=1}^{m_k} \bigcup \{\boldsymbol{\phi}^i(t_k)\}_{i=1}^{m}.$$

Note that: $t_1, t_k \in (\tau_1, \tau_\kappa]$; the set $\{\tau_n\}$ contains the high frequency IMU measurement time instants; and m_k are the total number of valid pseudorange measurements at time t_k.

Then, the objective is:

Objective 5.7. *Estimate the optimal state trajectory* $\mathbf{X} = [\boldsymbol{x}^\intercal(t_1), \ldots, \boldsymbol{x}^\intercal(t_K)]^\intercal \in \mathbb{R}^{n_s(K+1)}$ *and integers* $\mathbf{N} = [N^1, \ldots, N^m]^\intercal \in \mathbb{Z}^m$ *with the given sensor measurements* $\mathbf{U}$, $\mathbf{Y}$ *and the prior state density* $p_{\boldsymbol{x}}(\boldsymbol{x}(t_0))$. △
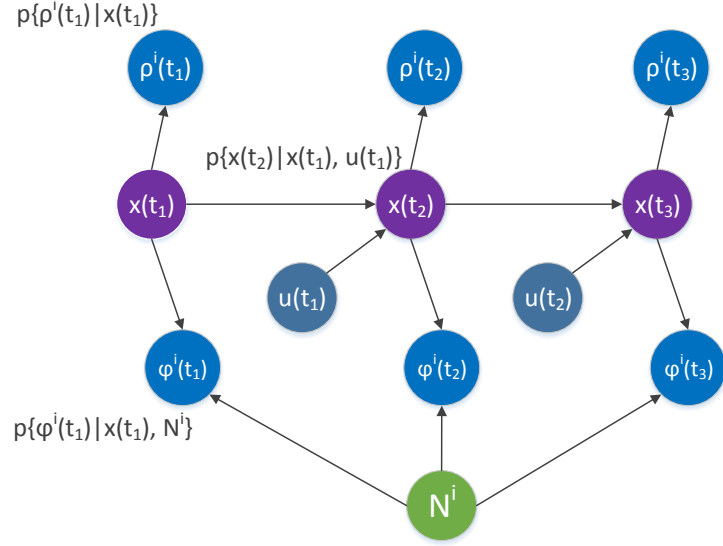
Figure 5.2: The RTK GPS/INS navigation estimation problem over a 3-epoch time window represented by a Bayes Newtwork. For clear illustration, only the GPS measurements from the i-th satellite are shown in the figure.

Fig. 5.2 represents the estimation problem summarized above by a Bayes Network (BN). For simplified illustration purpose, the GPS measurements from only the i-th satellite are shown in Fig. 5.2. This Bayes Network is a directed acyclic graph representing the joint distribution $p(\mathbf{X}, \mathbf{N}, \mathbf{Y}, \mathbf{U})$. Each node in the BN represents a variable and its associated conditional density. For example, in Fig. 5.2 the node labeled $\boldsymbol{x}(t_1)$ represents the conditional probability of the unknown state $\boldsymbol{x}(t_2)$ given $\boldsymbol{x}(t_1)$ and $\mathbf{U}_1$. So, $p(\mathbf{X}, \mathbf{N}, \mathbf{Y}, \mathbf{U})$ can be regarded as 'factorized' over the BN by following Bayes rule. This representation motivates another type of graphical model, named *Factor Graph* shown in Fig. 5.3. In a factor graph, only the unknown variables are explicitly labeled as nodes, while the edges (with black dots) represents

the 'factors' which encode the probabilistic information (constraints) between the unknown variables. In particular, in Fig. 5.3 the factor above the node $\boldsymbol{x}(t_1)$ is for the initial state prior. The factors between two consecutive purple nodes are the probabilistic constraints derived from the IMU data and kinematic constraints. The factors above the nodes $\{\boldsymbol{x}(t_k)\}_{k=1}^{3}$ are for the DGPS pseudorange measurements. The factors between the purple-green pairs are for the carrier phase measurements.
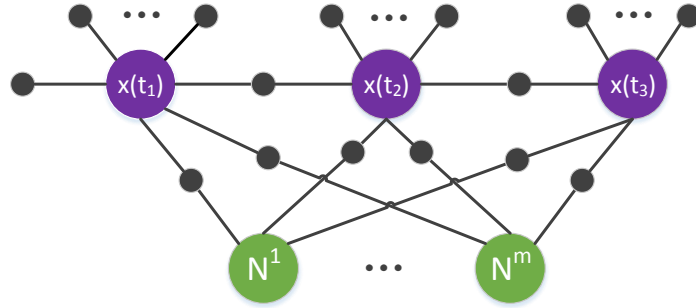
Figure 5.3: The RTK GPS/INS navigation estimation problem over a 3-second time window represented by a Factor Graph. The labeled nodes are for the unknown variables, and the black dotted edges are the factors encoding the probabilistic information between the unknown variables.

Let $\mathbf{X}_+ = \{\boldsymbol{x}(t) \text{ for } t = t_2, \ldots, t_K\}$, then the joint probability $p(\mathbf{X}, \mathbf{N}, \mathbf{Y}, \mathbf{U})$ can be factored as

$$p(\mathbf{X}, \mathbf{N}, \mathbf{Y}, \mathbf{U})$$

$$= p(\mathbf{X}, \mathbf{U}, \mathbf{N}) p(\mathbf{Y}|\mathbf{X}, \mathbf{U}, \mathbf{N})$$

$$= p(\mathbf{X}_+, \boldsymbol{x}(t_1), \mathbf{U}) p(\mathbf{Y}|\mathbf{X}, \mathbf{N})$$

$$= p(\boldsymbol{x}(t_1), \mathbf{U}) p(\mathbf{X}_+|\boldsymbol{x}(t_1), \mathbf{U}) p(\mathbf{Y}|\mathbf{X}, \mathbf{N})$$

$$= p(\boldsymbol{x}(t_1)) p(\mathbf{X}_+|\boldsymbol{x}(t_1), \mathbf{U}) p(\mathbf{Y}|\mathbf{X}, \mathbf{N}) \tag{5.27}$$

Thus, a *Maximum-a-Posteriori* problem can be formed to estimate $\mathbf{X}$ and $\mathbf{N}$ by maximizing the right hand side of eqn. (6.2),

$$\max_{\mathbf{X} \in \mathbb{R}^{n_s K}, \mathbf{N} \in \mathbb{Z}^m} p(\boldsymbol{x}(t_1)) p(\mathbf{X}_+|\boldsymbol{x}(t_1), \mathbf{U}) p(\mathbf{Y}|\mathbf{X}, \mathbf{N}). \tag{5.28}$$

With the Gaussian noise assumption, the negative log-likelihood of the right hand side of eqn. (6.2) is

$$\|\mathbf{v}(\mathbf{X}, \mathbf{N})\|_{\mathbf{W}}^2 = \|\boldsymbol{x}(t_1) - \boldsymbol{x}_1\|_{\mathbf{P}_1}^2$$

$$+ \sum_k \|\boldsymbol{\phi}(\boldsymbol{x}(t_k), \mathbf{U}_k) - \boldsymbol{x}(t_{k+1})\|_{\mathbf{Q}_k}^2$$

$$+ \sum_k \sum_i \|h_k^i(\boldsymbol{x}(t_k)) - \rho^i(t_k)\|_{\sigma_\rho^2}^2 \tag{5.29}$$

$$+ \sum_k \sum_i \|h_k^i(\boldsymbol{x}(t_k)) + \lambda N^i - \varphi^i(t_k)\|_{\sigma_\varphi^2}^2$$

where $\|\mathbf{v}\|_{\mathbf{W}}^2 = \mathbf{v}^\top \mathbf{W}^{-1} \mathbf{v}$ is the squared Mahalanobis distance with the matrix $\mathbf{W}$. All terms on the right-hand side also use this notation. The vector $\mathbf{v}$ is the concatenation

of each of the vectors summed in the right-hand side of eqn. (6.4). In eqn. (6.4), the operator ϕ and the covariance matrix $\mathbf{Q}_k$ are defined in Appendix II. The matrix $\mathbf{W}$ is the positive definite block diagonal matrix formed by the positive definite submatrices $\mathbf{Q}_k$, $\mathbf{P}_1$, σ_ρ^2 and σ_φ^2. Using MATLAB syntax, $\mathbf{W}$ could be represented as

$$\mathbf{W} = \mathrm{blkdiag}\big(\mathbf{P}_1, \mathbf{Q}_0, \ldots, \mathbf{Q}_{K-1},$$

$$\sigma_\rho^2, \ldots, \sigma_\rho^2, \sigma_\varphi^2, \ldots, \sigma_\varphi^2\big).$$

Recall that N^i is the integer for satellite i, which is assumed constant over the time interval, and $\rho^i(t_k)$ and $\varphi^i(t_k)$ are the code and phase measurements from satellite i at epoch t_k. Due to Assumption 6.1, the notation N^i does not require a time index. The IMU data set $\mathbf{U}_k$ contains all IMU measurements in the interval $[t_k, t_{k+1}]$.

Let $\mathbf{\Sigma}_W^\mathsf{T} \mathbf{\Sigma}_W = \mathbf{W}^{-1}$, then

$$\boldsymbol{r} \triangleq \mathbf{\Sigma}_W \mathbf{v} \tag{5.30}$$

is the weighted residual and $\|\mathbf{v}\|_{\mathbf{W}}^2 = \|\boldsymbol{r}\|^2$. For notation simplicity, herein we denote the tuple $(\mathbf{X}, \mathbf{N}) = [\mathbf{X}^\mathsf{T}, \mathbf{N}^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^{n_s K} \times \mathbb{Z}^m$. So, the MAP estimation is transformed into a Nonlinear Mixed Integer Least Square (NMILS) problem,

$$(\mathbf{X}^*, \mathbf{N}^*) = \underset{\mathbf{X} \in \mathbb{R}^{n_s K}, \mathbf{N} \in \mathbb{Z}^m}{arg\,min} \|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2, \tag{5.31}$$

where $\boldsymbol{r}$ is a vector representing the factors in the graph model:

$$\boldsymbol{r}(\mathbf{X}, \mathbf{N}) = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{P}_1} \left(\boldsymbol{x}(t_1) - \boldsymbol{x}_1 \right) \\[8pt] \boldsymbol{\Sigma}_{\mathbf{Q}_1} \left(\boldsymbol{\phi} \left(\boldsymbol{x}(t_1), \mathbf{U}_1 \right) - \boldsymbol{x}(t_2) \right) \\[8pt] \vdots \\[8pt] \boldsymbol{\Sigma}_{\mathbf{Q}_{K-1}} \left(\boldsymbol{\phi} \left(\boldsymbol{x}(t_{K-1}), \mathbf{U}_{K-1} \right) - \boldsymbol{x}(t_K) \right) \\[8pt] \sigma_\rho^{-1} \left(h_1^1 \left(\boldsymbol{x}(t_1) \right) - \rho^1(t_1) \right) \\[8pt] \vdots \\[8pt] \sigma_\rho^{-1} \left(h_K^m \left(\boldsymbol{x}(t_K) \right) - \rho^m(t_K) \right) \\[8pt] \sigma_\varphi^{-1} \left(h_1^1 \left(\boldsymbol{x}(t_1) \right) + \lambda N^1 - \varphi^1(t_K) \right) \\[8pt] \vdots \\[8pt] \sigma_\varphi^{-1} \left(h_K^m \left(\boldsymbol{x}(t_K) \right) + \lambda N^m - \varphi^m(t_K) \right) \end{bmatrix} . \tag{5.32}$$

5.4   Graph Optimization for RTK GPS/INS

The original CRT integer ambiguity resolution and trajectory estimation approach is proposed in [13], which can be outlined as the following three steps:

1. Obtain the *float solution* by neglecting the integeral nature of the ambiguity $\mathbf{N}$

$$(\check{\mathbf{X}}, \check{\mathbf{N}}) = \underset{(\mathbf{X}, \mathbf{N}) \in \mathbb{R}^{n_s K + m}}{arg\,min} \| \boldsymbol{r}(\mathbf{X}, \mathbf{N}) \|^2; \tag{5.33}$$

In this step, outlier rejection procedures presented in Chapter 4 can be executed;

2. Based on $(\check{\mathbf{X}}, \check{\mathbf{N}})$, solve Nonlinear Mixed Integer Least Square in eqn. (6.6) to get the optimal solution $(\mathbf{X}^*, \mathbf{N}^*)$;

3. Check the validity of integer estimates with integer validation techniques [77].

The above Step (1) can be solved by the robust optimization method proposed in Section 4.4.2. The aim of Step (2) is to deal with outlier measurements (if detected) and provide good initialization points for the NMILS in Step (2). In the following, the second step above of solving Nonlinear Mixed Integer Least Square (NMILS) is focused on.

Proposition 5.4 implies several ways to solve the NMILS. Herein, the iterative MILS method is presented as an example. To solve the optimization in eqn. (6.6) in an iterative manner, the *residual* $\boldsymbol{r}(\mathbf{X}, \mathbf{N})$ is linearized around the current estimates $(\hat{\mathbf{X}}, \hat{\mathbf{N}})$ as

$$\boldsymbol{r}(\mathbf{X}, \mathbf{N}) \approx \boldsymbol{r}(\hat{\mathbf{X}}, \hat{\mathbf{N}}) + \mathbf{J}(\hat{\mathbf{X}}, \hat{\mathbf{N}})(\delta\mathbf{X}, \delta\mathbf{N}), \tag{5.34}$$

where $\mathbf{J}(\hat{\mathbf{X}}, \hat{\mathbf{N}})$ is the Jacobian matrix of $\boldsymbol{r}(\mathbf{X}, \mathbf{N})$ evaluated at $(\hat{\mathbf{X}}, \hat{\mathbf{N}})$, and $(\delta\mathbf{X}, \delta\mathbf{N}) = (\mathbf{X}, \mathbf{N}) - (\hat{\mathbf{X}}, \hat{\mathbf{N}})$ is the estimation error. Furthermore, $\mathbf{J}(\hat{\mathbf{X}}, \hat{\mathbf{N}})$ can be decomposed as

$$\mathbf{J}(\hat{\mathbf{X}}, \hat{\mathbf{N}}) = [\boldsymbol{A}, \boldsymbol{B}],$$

where $\boldsymbol{A}$ contains the columns of $\mathbf{J}(\hat{\mathbf{X}}, \hat{\mathbf{N}})$ that are the partial with respect to $\mathbf{X}$ and $\boldsymbol{B}$ contains the partial with respect to $\mathbf{N}$. Thus, eqn. (5.34) can be rewritten as

$$\boldsymbol{r}(\mathbf{X}, \mathbf{N}) \approx \boldsymbol{r}(\hat{\mathbf{X}}, \hat{\mathbf{N}}) + \boldsymbol{A}\delta\mathbf{X} + \boldsymbol{B}\delta\mathbf{N}.$$

Then, the next step is to solve the following Mixed Integer Least Square (MILES) [11]

$$\min_{\delta\mathbf{X}\in\mathbb{R}^{n_sK},\delta\mathbf{N}\in\mathbb{Z}^m}\|\boldsymbol{r}(\hat{\mathbf{X}},\hat{\mathbf{N}})+\boldsymbol{A}\delta\mathbf{X}+\boldsymbol{B}\delta\mathbf{N}\|^2. \tag{5.35}$$

By dropping the notation $(\hat{\mathbf{X}},\hat{\mathbf{N}})$ in $\boldsymbol{r}(\hat{\mathbf{X}},\hat{\mathbf{N}})$ and defining the QR-decomposition

$$\boldsymbol{A}=[\boldsymbol{Q}_A,\bar{\boldsymbol{Q}}_A]\begin{bmatrix}\boldsymbol{R}_A\\\\\mathbf{0}\end{bmatrix}, \tag{5.36}$$

the cost function in eqn. (5.35) can be factorized as

$$\begin{aligned}&\|\boldsymbol{r}+\boldsymbol{A}\delta\mathbf{X}+\boldsymbol{B}\delta\mathbf{N}\|^2\\&=\left\|\begin{bmatrix}\boldsymbol{Q}_A^\intercal\\\\\bar{\boldsymbol{Q}}_A^\intercal\end{bmatrix}\boldsymbol{r}+\begin{bmatrix}\boldsymbol{R}_A\\\\\mathbf{0}\end{bmatrix}\delta\mathbf{X}+\begin{bmatrix}\boldsymbol{Q}_A^\intercal\boldsymbol{B}\\\\\bar{\boldsymbol{Q}}_A^\intercal\boldsymbol{B}\end{bmatrix}\delta\mathbf{N}\right\|^2\\&=\|\boldsymbol{Q}_A^\intercal\boldsymbol{r}+\boldsymbol{R}_A\delta\mathbf{X}+\boldsymbol{Q}_A^\intercal\boldsymbol{B}\delta\mathbf{N}\|^2+\|\bar{\boldsymbol{Q}}_A^\intercal\boldsymbol{r}+\bar{\boldsymbol{Q}}_A^\intercal\boldsymbol{B}\delta\mathbf{N}\|^2.\end{aligned} \tag{5.37}$$

Note that for any fixed $\delta\mathbf{N}$, the first term on the right hand side of the above equation can be made equal to zero by appropriate choice of $\delta\mathbf{X}$. Thus, to solve the following Integer Least Square (ILS) will lead to the optimum of eqn. (5.35),

$$\min_{\delta\mathbf{N}\in\mathbb{Z}^m}\|\bar{\boldsymbol{Q}}_A^\intercal\boldsymbol{r}+\bar{\boldsymbol{Q}}_A^\intercal\boldsymbol{B}\delta\mathbf{N}\|^2. \tag{5.38}$$

The standard ILS method is reviewed in Section 5.1.4.

5.5 Implementation Results

This section presents implementation results of the graph optimization based RTK GPS/INS.To evaluate performance the graph optimization based RTK GPS/INS,

the proposed approach is implemented to post-process several GPS/INS data sets collected from an automotive vehicle. The vehicle was equipped with a 200Hz MEMS IMU and a NovAtel OEMV3 receiver. The differential corrections are from the UC Riverside Ntrip caster (ntrip.engr.ucr.edu:2101) which broadcasts raw dual frequency GPS measurements publicly over the internet in 1Hz. The baseline from rover to base is within 10km. This presentation will consider two data sets: 1008-second stationary data set and a 340-second moving data set. Besides of L1 measurements, wide-lane phase measurements are formed to facilitate the integer ambiguity resolution. For test purpose, all m-second CRT windows sequentially picked from the data set are tested independently, without any prior knowledge from previous estimation steps. In practical implementations, sliding window [20, 82] or incremental scheme [35, 58] can be applied to utilize the prior knowledge got from previous estimation steps. Furthermore, the integer ambiguity resolution may not need executing for every CRT window. If the integer estimates are validated and no cycle slip happens, the integer estimates can be fixed for later use.

Stationary Data

In this implementation with 1008-seconds of stationary data set, a 3-second CRT window is used for integer ambiguity resolution. The algorithm itself has no knowledge that the system is stationary. In particular, for each 3-second CRT windows, $\mathbf{X} = [\boldsymbol{x}(t_0)^\intercal, \boldsymbol{x}(t_1)^\intercal, \boldsymbol{x}(t_2)^\intercal, \boldsymbol{x}(t_3)^\intercal]^\intercal$ is estimated. The ground truth of the parking
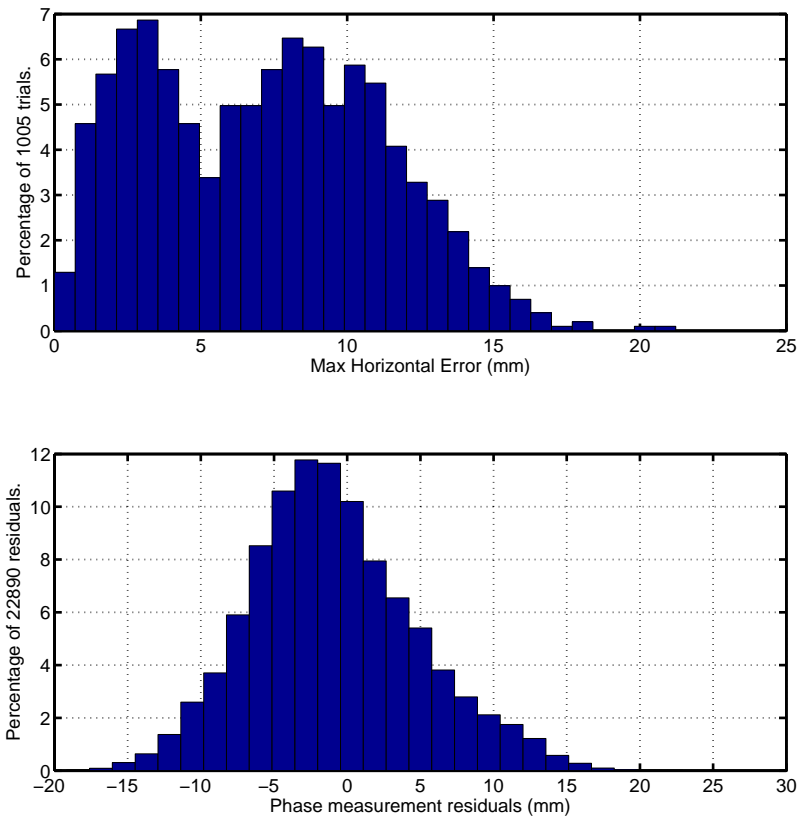
Figure 5.4: Implementation results with 1008 seconds of stationary data. The number of available satellites is 7-9. Window length is 3 seconds. Trials are independent with zero initials. All 3-second windows in this 1008-second data set are tested with the proposed CRT approach.

position is surveyed with millimeter accuracy. To compute the error, the position estimation result is down-sampled to 10Hz and compared with the ground truth. Fig. 5.4 shows the statistics of the average horizontal positioning errors (in millimeters) of each trial. The number of available satellites (i.e., m in Assumption 1) throughout this data set is 7-9. Fig. 5.4 shows the positioning errors of 1005 trials are mostly below 1 centimeter which matches the RTK accuracy expectation. Thus, it can be claimed that the successful rate of integer ambiguity resolution of the proposed method with this 1008 second stationary data set is 100% with 3-second CRT window. The totally 22890 estimated L1 residuals of carrier phase measurements fit a normal distribution with $\mu = -0.000941$ and $\sigma = 0.0058$ in meters, showing the performance of the estimator.

Moving Data

In this implementation with a 340-second moving data set collected from a urban area with $72km/h$ speed limit, a 5-second CRT window is used for integer ambiguity resolution. Fig. 5.5 shows the trajectory on Google Map where the yellow mark is the starting point and the blue mark is the ending point. To ensure good constellation geometry to solve integers, only the time intervals with 7 or more satellites are used for algorithm tests. To reduce outliers, a 5° elevation mask was applied while only the highest 8 satellites (at most) are used for estimation. In the tests, only the results of validated integer estimation were recorded. As a result, totally 290 trials

Figure 5.5: Trajectory of the 340s moving data on Google Map.

are available to solve and validate integer successfully by picking different intervals from the data set. For performance evaluation, the 340-second was processed by the full-batch approach presented in [78] to provide ground truth. The upper plot of Fig. 5.6 shows that most of the average horizontal errors are below 5 centimeters. The lower plot of Fig. 5.6 shows that the estimated residuals are at the centimeter level. Notice that the tails length of the residuals distribution shown in Fig. 5.6 is determined by the significant level α or other thresholds picked by the designer to do outlier rejections through hypothesis tests.
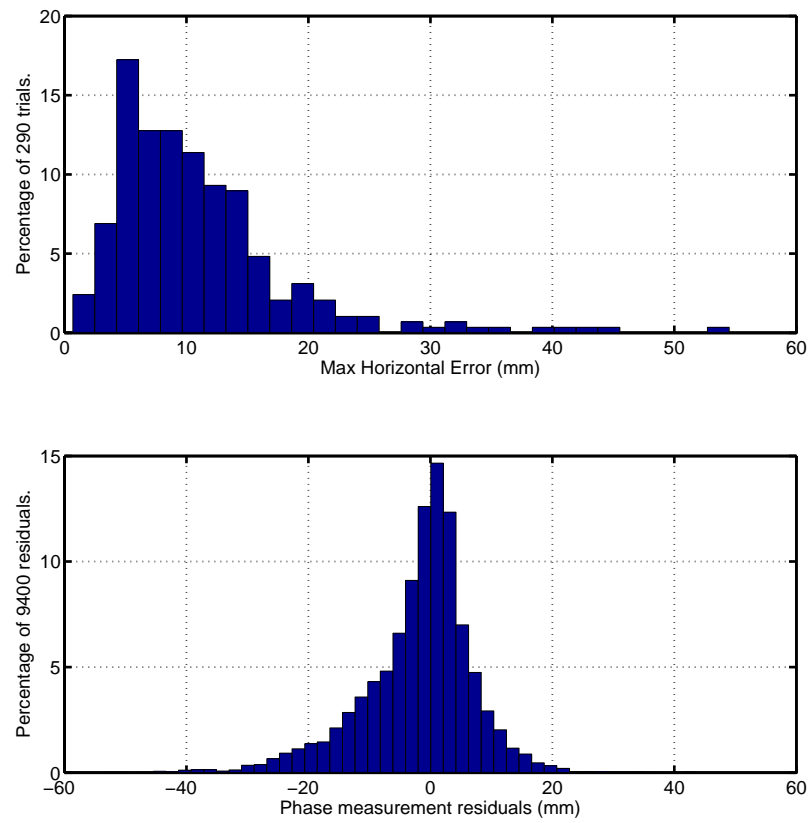
Figure 5.6: Implementation results with 340 seconds of moving data. Only the intervals with the number of available satellites between 7-8 are used for trials. Window length is 5 seconds. Totally 290 trials solved and validated integer with the proposed CRT method.

Chapter 6

Computationally Efficient Integer Ambiguity Resolution: a Common-Position-Shift method

In Chapter 5, although increasing the length of the CRT window would improve the estimation accuracy, the corresponding high computation leads to challenges of implementing the CRT integer ambiguity resolution in real-time, especially on resource limited platforms. Since the time complexity of standard QR-decomposition is $2mn^2$ (flops) assuming A is a $m \times n$ matrix, the eqn. (5.36) takes a large portion of total computation in integer ambiguity resolution. On the other hand, the computation of the ILS in eqn. (5.38) will not vary significantly, since the dimension of $\delta \mathbf{N}$

is upper-bounded by 15 (for each carrier frequency) in practice.[1] Also, to evaluate the INS residuals (see Appendix I), the integration process with high frequency IMU measurements also take a considerable portion of the total computation.

The main contribution of this dissertation is to propose an alternative method to solve the NMILS represented in eqn. (6.6) with lower computational requirements. This improvement is realized through reasonable approximations which are valid for engineering purpose (i.e., the error introduced is at least a factor of ten smaller than the measurement noise). The proposed method is inspired by the observation in practice that the optimal solution trajectory $\mathbf{X}^*$ (with integers resolved) is only different from the float solution $\check{\mathbf{X}}$ in terms of a common 3D position error of each state in the trajectory (e.g., see Fig. 6.1). Mathematical analysis in Section 6.3 verifies this observation.

Mathematical analysis shows that an estimation of the common 3D position shift can replace the above Step (2) and achieves equivalent accuracy. The dimension of the problem is reduced significantly and the heavy INS reintegration for iterative relinearization is avoided. The procedures and analysis of this Common Position Shift (CPS) approach are described in the following sections.

---

[1]The upper bound 15 for each carrier frequency is derived from the practical availability of GPS satellites during a feasible time interval.
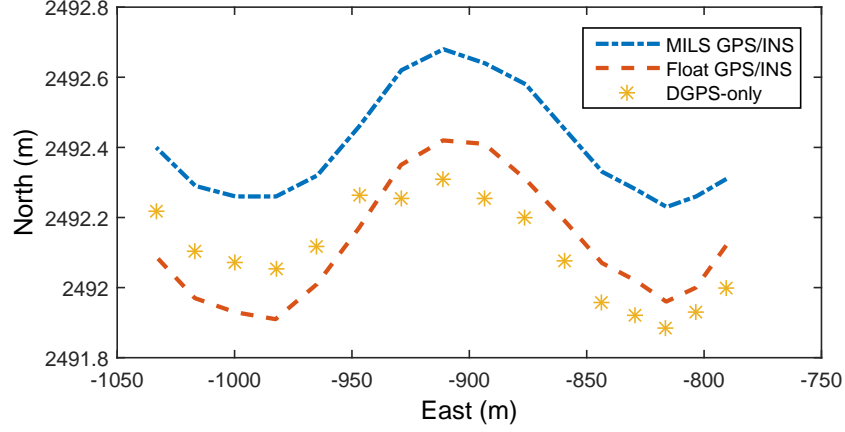
Figure 6.1: An example of horizontal positioning results comparison. The dot-dashed (blue) curve is for the RTK GPS/INS (MILS GPS/INS) solution in Chapter 5 which has centimeter level global positioning accuracy. The dashed (red) curve is for the float solution in eqn. (5.33). The asterisks (yellow) denotes the DGPS positioning without integrated with INS.

6.1   Problem Statement

This chapter also considers the smoothing problem of RTK GPS/INS over a time window. Most of notation used in this chapter can be found in Chapter 5. Herein, several extra definitions and assumptions are made to facilitate the presentation of the proposed approach in this chapter. For the convenience of later discussion, denote

$$\boldsymbol{s}(t) = [\boldsymbol{v}^{\mathsf{T}}(t), \boldsymbol{q}^{\mathsf{T}}(t), \boldsymbol{b}_a^{\mathsf{T}}(t), \boldsymbol{b}_g^{\mathsf{T}}(t)]^{\mathsf{T}} \in \mathbb{R}^{n_s - 3} \tag{6.1}$$

as the vector of system states except 3D position, and then we have $\boldsymbol{x}(t) = [\boldsymbol{p}^{\mathsf{T}}(t), \boldsymbol{s}^{\mathsf{T}}(t)]^{\mathsf{T}}$. Similarly, $\boldsymbol{x}_1 = [\boldsymbol{p}_1^{\mathsf{T}}, \boldsymbol{s}_1^{\mathsf{T}}]^{\mathsf{T}}$ is for the prior. Similar notations also apply to those of state estimates $\hat{\boldsymbol{x}}(t) = [\hat{\boldsymbol{p}}^{\mathsf{T}}(t), \hat{\boldsymbol{s}}^{\mathsf{T}}(t)]^{\mathsf{T}}$.

For the simplicity of discussion in this paper, the following assumptions are made.

Assumption 6.1. *For $t \in [t_1, t_K]$, the receiver provides valid carrier phase measurements for m satellites, without loss of lock.* △

Assumption 6.2. *There is only prior distribution $(\boldsymbol{s}_1, \mathbf{P}_{s_1})$ for $\boldsymbol{s}(t_1)$. The position estimate $\hat{\boldsymbol{p}}(t_1)$ can be initialized with the GPS measurements at t_1.* △

Assumption 6.3. *The frequency of GPS measurements is $1Hz$.* △

With Assumption 6.1, the unknown integers in the carrier phase measurements from these m satellites are constants over $[t_1, t_K]$. Assumption 2 indicates the practice that the rover starts around an unknown location and the position estimate in the navigation system is initialized with the first GPS positioning result. Other state estimates $\hat{\boldsymbol{s}}(t_1)$ (velocity, attitude, biases) can be initialized with the prior $(\boldsymbol{s}_1, \mathbf{P}_{s_1})$ from other sources. Since in later analysis, the sampling rate of GPS measurements is used, so herein Assumption 3 is made for the simplicity of presenting the main idea of this paper. The proposed method can be generalized with other measurement frequency straightforwardly.

Under Assumption 6.1-3, the CRT estimation problem considered in this paper can be stated as follows.

For a system described by eqn. (3.1), we have

- an initial distribution for the state $\boldsymbol{s}(t_1) \sim \mathcal{N}(\boldsymbol{s}_1, \mathbf{P}_{s_1})$,

- IMU measurements $\mathbf{U} = \{\mathbf{U}_k\}_{k=1}^{K-1}$, where

$$\mathbf{U}_k = \{\tilde{\boldsymbol{u}}(\tau_n), t_k \leq \tau_n \leq t_{k+1}\},$$

- DGPS code and carrier phase measurements $\mathbf{Y} = \{\mathbf{Y}_k\}_{k=1}^{K}$, where

$$\mathbf{Y}_k = \{\rho^i(t_k)\}_{i=1}^{m_k} \bigcup \{\varphi^i(t_k)\}_{i=1}^{m}.$$

Note that: t_1, $t_k \in (\tau_1, \tau_\kappa]$; the set $\{\tau_n\}$ contains the high frequency IMU measurement time instants; and m_k are the total number of valid pseudorange measurements at time t_k. For simplicity of discussion in this paper, it is assumed that $m_k \equiv m$. The method presented in this paper can be extended to more complicated cases straightforwardly.

Then, the objective is:

Objective 6.4. *Estimate the optimal state trajectory* $\mathbf{X} \triangleq [\boldsymbol{x}^\intercal(t_1), \ldots, \boldsymbol{x}^\intercal(t_K)]^\intercal \in \mathbb{R}^{Kn_s}$ *and integers* $\mathbf{N} \triangleq [N^1, \ldots, N^m]^\intercal \in \mathbb{Z}^m$ *with the given sensor measurements* $\mathbf{U}$, $\mathbf{Y}$ *and the prior state density* $p_{\boldsymbol{s}}(\boldsymbol{s}(t_1))$. $\triangle$

In Chapter 5, the above objective is achieved by formulating and solving the corresponding *Maximum-a-Posteriori* (MAP) estimation problem. The accuracy and reliability of the solution is achieved by the proposed Nonlinear Mixed Integer Least Square method and faulty data removal scheme. Since the considered problem in this chapter is a bit different from that of Chapter 5, the derivation of the MAP problem is rerun as follows for completeness.

Let $\mathbf{X}_+ = \{\boldsymbol{x}(t) \text{ for } t = t_2, \ldots, t_K\}$, then the joint probability $p(\mathbf{X}, \mathbf{N}, \mathbf{Y}, \mathbf{U})$ can be factored as

$$p(\mathbf{X}, \mathbf{N}, \mathbf{Y}, \mathbf{U})$$

$$= p(\mathbf{X}, \mathbf{U}, \mathbf{N})p(\mathbf{Y}|\mathbf{X}, \mathbf{U}, \mathbf{N})$$

$$\propto p(\mathbf{X}_+, \boldsymbol{x}(t_1), \mathbf{U})p(\mathbf{Y}|\mathbf{X}, \mathbf{N})$$

$$= p(\boldsymbol{x}(t_1), \mathbf{U})p(\mathbf{X}_+|\boldsymbol{x}(t_1), \mathbf{U})p(\mathbf{Y}|\mathbf{X}, \mathbf{N})$$

$$= p(\boldsymbol{x}(t_1))p(\mathbf{X}_+|\boldsymbol{x}(t_1), \mathbf{U})p(\mathbf{Y}|\mathbf{X}, \mathbf{N})$$

$$\propto p(\boldsymbol{s}(t_1))p(\mathbf{X}_+|\boldsymbol{x}(t_1), \mathbf{U})p(\mathbf{Y}|\mathbf{X}, \mathbf{N}) \tag{6.2}$$

Thus, a *Maximum-a-Posteriori* problem can be formed to estimate $\mathbf{X}$ and $\mathbf{N}$ by maximizing the right hand side of eqn. (6.2),

$$\max_{\mathbf{X} \in \mathbb{R}^{n_s K}, \mathbf{N} \in \mathbb{Z}^m} p(\boldsymbol{s}(t_1))p(\mathbf{X}_+|\boldsymbol{x}(t_1), \mathbf{U})p(\mathbf{Y}|\mathbf{X}, \mathbf{N}). \tag{6.3}$$

With the Gaussian noise assumption, the negative log-likelihood of the right hand side of eqn. (6.2) is

$$\|\mathbf{v}(\mathbf{X}, \mathbf{N})\|_{\mathbf{W}}^2 = \|\boldsymbol{s}(t_1) - \boldsymbol{s}_1\|_{\mathbf{P}_{s_1}}^2$$

$$+ \sum_k \|\boldsymbol{\phi}(\boldsymbol{x}(t_k), \mathbf{U}_k) - \boldsymbol{x}(t_{k+1})\|_{\mathbf{Q}_k}^2$$

$$+ \sum_k \sum_i \|h_k^i(\boldsymbol{x}(t_k)) - \rho^i(t_k)\|_{\sigma_\rho^2}^2 \tag{6.4}$$

$$+ \sum_k \sum_i \|h_k^i(\boldsymbol{x}(t_k)) + \lambda N^i - \varphi^i(t_k)\|_{\sigma_\varphi^2}^2$$

where $\|\mathbf{v}\|_{\mathbf{W}}^2 = \mathbf{v}^\top \mathbf{W}^{-1} \mathbf{v}$ is the squared Mahalanobis distance with the matrix $\mathbf{W}$. All terms on the right-hand side also use this notation. The vector $\mathbf{v}$ is the concatenation

of each of the vectors summed in the right-hand side of eqn. (6.4). In eqn. (6.4), the operator ϕ and the covariance matrix $\mathbf{Q}_k$ are defined in Appendix II. The matrix $\mathbf{W}$ is the positive definite block diagonal matrix formed by the positive definite submatrices $\mathbf{Q}_k$, $\mathbf{P}_{s_1}$, σ_ρ^2 and σ_φ^2. Using MATLAB syntax, $\mathbf{W}$ could be represented as

$$\mathbf{W} = \text{blkdiag}\big(\mathbf{P}_{s_1}, \mathbf{Q}_0, \ldots, \mathbf{Q}_{K-1},$$

$$\sigma_\rho^2, \ldots, \sigma_\rho^2, \sigma_\varphi^2, \ldots, \sigma_\varphi^2\big).$$

Recall that N^i is the integer for satellite i, which is assumed constant over the time interval, and $\rho^i(t_k)$ and $\varphi^i(t_k)$ are the code and phase measurements from satellite i at epoch t_k. Due to Assumption 6.1, the notation N^i does not require a time index. The IMU data set $\mathbf{U}_k$ contains all IMU measurements in the interval $[t_k, t_{k+1}]$.

Let $\boldsymbol{\Sigma}_W^\mathsf{T} \boldsymbol{\Sigma}_W = \mathbf{W}^{-1}$, then

$$\boldsymbol{r} \triangleq \boldsymbol{\Sigma}_W \mathbf{v} \tag{6.5}$$

is the weighted residual and $\|\mathbf{v}\|_{\mathbf{W}}^2 = \|\boldsymbol{r}\|^2$. For notation simplicity, herein we denote the tuple $(\mathbf{X}, \mathbf{N}) = [\mathbf{X}^\mathsf{T}, \mathbf{N}^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^{n_s K} \times \mathbb{Z}^m$. So, the MAP estimation is transformed into a Nonlinear Mixed Integer Least Square (NMILS) problem,

$$(\mathbf{X}^*, \mathbf{N}^*) = \underset{\mathbf{X} \in \mathbb{R}^{n_s K}, \mathbf{N} \in \mathbb{Z}^m}{arg\,min} \|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2, \tag{6.6}$$

where $\boldsymbol{r}$ is a vector representing the factors in the graph model:

$$
\boldsymbol{r}(\mathbf{X}, \mathbf{N}) =
\begin{bmatrix}
\boldsymbol{\Sigma}_{\mathbf{P}_{s_1}} \left(\boldsymbol{s}(t_1) - \boldsymbol{s}_1\right) \\[6pt]
\boldsymbol{\Sigma}_{\mathbf{Q}_1} \left(\boldsymbol{\phi}\big(\boldsymbol{x}(t_1), \mathbf{U}_1\big) - \boldsymbol{x}(t_2)\right) \\[6pt]
\vdots \\[6pt]
\boldsymbol{\Sigma}_{\mathbf{Q}_{K-1}} \left(\boldsymbol{\phi}\big(\boldsymbol{x}(t_{K-1}), \mathbf{U}_{K-1}\big) - \boldsymbol{x}(t_K)\right) \\[6pt]
\sigma_\rho^{-1} \left(h_1^1\big(\boldsymbol{x}(t_1)\big) - \rho^1(t_1)\right) \\[6pt]
\vdots \\[6pt]
\sigma_\rho^{-1} \left(h_K^m\big(\boldsymbol{x}(t_K)\big) - \rho^m(t_K)\right) \\[6pt]
\sigma_\varphi^{-1} \left(h_1^1\big(\boldsymbol{x}(t_1)\big) + \lambda N^1 - \varphi^1(t_K)\right) \\[6pt]
\vdots \\[6pt]
\sigma_\varphi^{-1} \left(h_K^m\big(\boldsymbol{x}(t_K)\big) + \lambda N^m - \varphi^m(t_K)\right)
\end{bmatrix}. \tag{6.7}
$$

6.2   Common Position Shift Estimation

This section presents the proposed Common Position Shift (CPS) method which is an computationally efficient alternative to the original CRT method proposed in Section 5.4. The key point is to construct a smaller optimization to replace that in Step (2) of Section 5.4. First, the notation for the CPS method is defined.
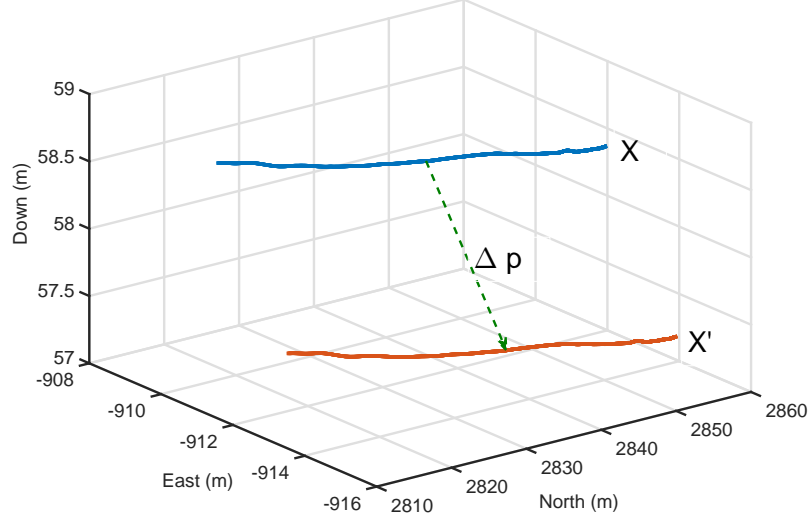
Figure 6.2: An example of 3D position shift: the blue curve is the original trajectory, the green dashed arrow represents the 3D position shift vector and the red curve is the shifted trajectory.

6.2.1 Notation of Common-Position-Shift

Given an estimated trajectory $\mathbf{X} = [\boldsymbol{x}^{\mathsf{T}}(t_1), \ldots, \boldsymbol{x}^{\mathsf{T}}(t_K)]^{\mathsf{T}}$ and a common-position-shift vector $\Delta\boldsymbol{p} \in \mathbb{R}^3$, define the *Common Position Shift* operator $\oplus$ as

$$\mathbf{X}' = \mathbf{X} \oplus \Delta\boldsymbol{p} \triangleq [\boldsymbol{x}^{\mathsf{T}}(t_1) \oplus \Delta\boldsymbol{p}, \ldots, \boldsymbol{x}^{\mathsf{T}}(t_K) \oplus \Delta\boldsymbol{p}]^{\mathsf{T}}$$

which denotes adding the constant vector $\Delta\boldsymbol{p}$ to the position estimate portion, $\boldsymbol{p}(t_k)$, of each state estimate $\boldsymbol{x}(t_k)$ in $\mathbf{X}$. The resulting trajectory $\mathbf{X}'$ is referred as the *shifted* trajectory with respect to the original $\mathbf{X}$. Fig. 6.2 gives an example of the 3D common-position-shift.

140

6.2.2 Outline of the CPS method

This section gives an outline of the proposed CPS method. First, some notation related to the cost functions is introduced.

We will show that cost function $\|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2$ of the NMILS problem (6.6) can be rewritten as a sum of cost functions:

$$\|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2 = \|\boldsymbol{r}_1(\mathbf{X})\|^2 + \|\boldsymbol{r}_2(\mathbf{X})\|^2 + \|\boldsymbol{r}_3(\mathbf{X}, \mathbf{N})\|^2. \tag{6.8}$$

We will also show two important related properties. First, the term $\|\boldsymbol{r}_1(\mathbf{X}, \mathbf{N})\|^2$ determines the shape and general location of the trajectory, but is insensitive to a common position shift $\Delta\boldsymbol{p}$ and to the integer vector $\mathbf{N}$. Second, that $\|\boldsymbol{r}_2(\mathbf{X})\|^2 + \|\boldsymbol{r}_3(\mathbf{X}, \mathbf{N})\|^2$ can be minimized solely by the choice of $(\Delta\boldsymbol{p}, \mathbf{N})$. Therefore, if the linearization errors are ignored, Proposition 6.10 will show that the cost function can be rewritten as

$$\|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2 = \|\boldsymbol{r}_1(\check{\mathbf{X}})\|^2 + \|\boldsymbol{r}_2(\check{\mathbf{X}} \oplus \Delta\boldsymbol{p})\|^2 + \|\boldsymbol{r}_3(\check{\mathbf{X}} \oplus \Delta\boldsymbol{p}, \mathbf{N})\|^2$$

where $\mathbf{X} = \check{\mathbf{X}} \oplus \Delta\boldsymbol{p}$.

These facts allow us to redesign the algorithm of [13] that is summarized in Section 5.4 to the following:

1. Find either the float solution $\check{\mathbf{X}}$ defined in eqn. (5.33) or the integer free solution $\mathbf{X}^\circledast$ defined in eqn. (6.15), which are shown to be identical in Proposition 6.5.

2. Find $(\Delta\boldsymbol{p}^\star, \mathbf{N}^\star)$ that is the optimal solution of

$$\min_{\Delta\boldsymbol{p}\in\mathbb{R}^3, \mathbf{N}\in\mathbb{Z}^m} \|\boldsymbol{r}_2(\check{\mathbf{X}}\oplus\Delta\boldsymbol{p})\|^2 + \|\boldsymbol{r}_3(\check{\mathbf{X}}\oplus\Delta\boldsymbol{p}, \mathbf{N})\|^2, \tag{6.9}$$

where $\check{\mathbf{X}}$ is fixed when evaluating $\|\boldsymbol{r}_2\|^2 + \|\boldsymbol{r}_3\|^2$.

3. Check the validity of the integer estimates.

This cost function $\|\boldsymbol{r}_2\|^2 + \|\boldsymbol{r}_3\|^2$ is derived through processing the double difference GPS pseudorange and carrier phase measurements. The details of the mathematical analysis are described in the remainder of the article.

The Common Position Shift (CPS) estimation in eqn. (6.9) is designed to replace the original full NMILS Step (2) in Section 5.4, Note that $\check{\mathbf{X}}$ is the float solution which is considered constant when evaluating the cost $\|\boldsymbol{r}_2\|^2$. The optimization in eqn. (6.9) is also solved by the NMILS method outlined in Section 5.4. The integer estimates are initialized by rounding the float estimates $\check{\mathbf{N}}$ provided by either solution approach of Proposition 6.5. Compared with solving the full NMILS in eqn. (6.6) directly, the computational cost in eqn. (6.9) is significantly reduced due to the much smaller dimension of the real unknown variable $\Delta\boldsymbol{p} \in \mathbb{R}^3$ versus $\mathbf{X} \in \mathbb{R}^{n_s K}$. In particular, this dimension reduction facilitates the QR-decomposition in eqn. (5.36). Thus, for our 6D RTK GPS/INS integer ambiguity resolution problem summarized in Section ??, the dimension of corresponding A matrices in full MILS and CPS MILS are $M \times N_s$ versus $2m \times 3$, where $M = n_s K + 2mK - 3$ is the total measurement number, $N_s = n_s K$ is the total state dimension and $2m$ is the total GPS measurement number

(code and carrier phase). Furthermore, in the CPS NMILS (6.9), INS reintegrations, which bring heavy computations, are not required any more.

6.2.3 Decomposition of GPS cost terms

This subsection shows how the decomposition in eqn. (6.8) applies to the GPS code and phase measurements. Define $\boldsymbol{\varphi}^i = [\varphi^i(t_1), \ldots, \varphi^i(t_K)]^\mathsf{T} \in \mathbb{R}^K$ to be the vector stacking the carrier phase measurements of the i-th satellite. The last summation term in eqn. (6.4) can be rewritten as

$$
\begin{aligned}
&\sum_k \sum_i \|h_k^i\big(\boldsymbol{x}(t_k)\big) + \lambda N^i - \varphi^i(t_k)\|_{\sigma_\varphi^2}^2 \\
&= \sum_i \left[\sum_k \|h_k^i\big(\boldsymbol{x}(t_k)\big) + \lambda N^i - \varphi^i(t_k)\|_{\sigma_\varphi^2}^2 \right] \\
&= \sum_i \|\boldsymbol{h}^i(\mathbf{X}) + \lambda \mathbf{1} N^i - \boldsymbol{\varphi}^i\|_{\sigma_\varphi^2 \mathbf{I}}^2,
\end{aligned}
\tag{6.10}
$$

where $\boldsymbol{h}^i = [h_1^i\big(\boldsymbol{x}(t_1)\big), \ldots, h_K^i\big(\boldsymbol{x}(t_K)\big)]^\mathsf{T} \in \mathbb{R}^K$, $\mathbf{1} = [1, \ldots, 1]^\mathsf{T} \in \mathbb{R}^K$, and $\mathbf{I}$ is the $K \times K$ identity matrix. In particular, $\mathbf{1}$ is rank 1 and can be QR-decomposed as

$$
[\mathcal{Q}_\mathbf{1}, \bar{\mathcal{Q}}_\mathbf{1}] \begin{bmatrix} \mathcal{R}_\mathbf{1} \\ \mathbf{0} \end{bmatrix} = \mathbf{1},
$$

where $[\mathcal{Q}_\mathbf{1}^\mathsf{T}]_{1 \times K}$ and $[\bar{\mathcal{Q}}_\mathbf{1}^\mathsf{T}]_{(K-1) \times K}$ are mappings to the column space and the left null space of $\mathbf{1}$. Note that we use $\mathcal{Q}_\mathbf{1}$ and $\mathbf{Q}_1$ as two different notations: the former represents the column spans the column space of $\mathbf{1}$ while the latter represent the covariance matrix of the first INS cost term in eqn. (6.4). Denote $\mathcal{Q} = [\mathcal{Q}_\mathbf{1}, \bar{\mathcal{Q}}_\mathbf{1}]$,

which is a unitary mapping. This QR-decomposition can be computed offline for different K.

The i-th summand in (6.10) can be decomposed into two parts through projecting it on the column space and the left null space of $\mathbf{1}$,

$$\|\boldsymbol{h}^i(\mathbf{X}) + \lambda\mathbf{1}N^i - \boldsymbol{\varphi}^i\|_{\sigma_\varphi^2\mathbf{I}}^2$$

$$= \left\|\mathcal{Q}^\mathsf{T}\left(\boldsymbol{h}^i(\mathbf{X}) + \lambda\mathbf{1}N^i - \boldsymbol{\varphi}^i\right)\right\|_{\mathcal{Q}^\mathsf{T}\sigma_\varphi^2\mathcal{Q}}^2$$

$$= \left\|\begin{bmatrix} \mathcal{Q}_1^\mathsf{T} \\ \bar{\mathcal{Q}}_1^\mathsf{T} \end{bmatrix}\left[\boldsymbol{h}^i(\mathbf{X}) + \lambda\mathcal{Q}\begin{bmatrix} \mathcal{R}_1 \\ \mathbf{0} \end{bmatrix}N^i - \boldsymbol{\varphi}^i\right]\right\|_{\sigma_\varphi^2\mathbf{I}}^2$$

$$= \left\|\begin{bmatrix} \mathcal{Q}_1^\mathsf{T}\boldsymbol{h}^i(\mathbf{X}) + \lambda\mathcal{R}_1 N^i - \mathcal{Q}_1^\mathsf{T}\boldsymbol{\varphi}^i \\ \bar{\mathcal{Q}}_1^\mathsf{T}\left(\boldsymbol{h}^i(\mathbf{X}) - \boldsymbol{\varphi}^i\right) \end{bmatrix}\right\|_{\sigma_\varphi^2\mathbf{I}}^2$$

$$= \left\|\bar{\mathcal{Q}}_1^\mathsf{T}\left(\boldsymbol{h}^i(\mathbf{X}) - \boldsymbol{\varphi}^i\right)\right\|_{\sigma_\varphi^2\mathbf{I}^-}^2$$

$$\quad + \left\|\mathcal{Q}_1^\mathsf{T}\boldsymbol{h}^i(\mathbf{X}) + \lambda\mathcal{R}_1 N^i - \mathcal{Q}_1^\mathsf{T}\boldsymbol{\varphi}^i\right\|_{\sigma_\varphi^2}^2, \tag{6.11}$$

where $\mathbf{I}^- = \bar{\mathcal{Q}}_1^\mathsf{T}\bar{\mathcal{Q}}_1$ is the $(K-1) \times (K-1)$ identity matrix. Note that the first term in eqn. (6.11) is independent of the integer ambiguity N^i.

Applying the same QR-factorization to the pseudorange summation term in eqn. (6.4) and reorganizing yields:

$$\|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2 = \|\boldsymbol{s}(t_1) - \boldsymbol{s}_1\|_{\mathbf{P}_{s_1}}^2 + \sum_k \|\boldsymbol{\phi}\big(\boldsymbol{x}(t_k), \mathbf{U}_k\big) - \boldsymbol{x}(t_{k+1})\|_{\mathbf{Q}_k}^2$$

$$+ \sum_i \left\|\bar{\mathcal{Q}}_1^\mathsf{T}\left(\boldsymbol{h}^i(\mathbf{X}) - \boldsymbol{\rho}^i\right)\right\|_{\sigma_\rho^2\mathbf{I}^-}^2 + \sum_i \left\|\bar{\mathcal{Q}}_1^\mathsf{T}\left(\boldsymbol{h}^i(\mathbf{X}) - \boldsymbol{\varphi}^i\right)\right\|_{\sigma_\varphi^2\mathbf{I}^-}^2$$

$$+ \sum_i \left\|\mathcal{Q}_1^\mathsf{T}\boldsymbol{h}^i(\mathbf{X}) - \mathcal{Q}_1^\mathsf{T}\boldsymbol{\rho}^i\right\|_{\sigma_\rho^2}^2 + \sum_i \left\|\mathcal{Q}_1^\mathsf{T}\boldsymbol{h}^i(\mathbf{X}) + \lambda\mathcal{R}_1 N^i - \mathcal{Q}_1^\mathsf{T}\boldsymbol{\varphi}^i\right\|_{\sigma_\varphi^2}^2.$$

Based on the above expression, for the presentation of the approach and its analysis, it will be convenient to define the following three cost functions. The first cost function:

$$\|\boldsymbol{r}_1(\mathbf{X})\|^2 \triangleq \|\boldsymbol{s}(t_1) - \boldsymbol{s}_1\|_{\mathbf{P}_{s1}}^2 + \sum_k \|\boldsymbol{\phi}\big(\boldsymbol{x}(t_k), \mathbf{U}_k\big) - \boldsymbol{x}(t_{k+1})\|_{\mathbf{Q}_k}^2$$
$$+ \sum_i \|\bar{\mathcal{Q}}_{\mathbf{1}}^{\mathsf{T}} \big(\boldsymbol{h}^i(\mathbf{X}) - \boldsymbol{\rho}^i\big)\|_{\sigma_\rho^2}^2 + \sum_i \|\bar{\mathcal{Q}}_{\mathbf{1}}^{\mathsf{T}} \big(\boldsymbol{h}^i(\mathbf{X}) - \boldsymbol{\varphi}^i\big)\|_{\sigma_\varphi^2 \mathbf{I}^-}^2$$

neglects the last two terms in $\|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2$. The second cost function

$$\|\boldsymbol{r}_2(\mathbf{X})\|^2 \triangleq \sum_i \|\mathcal{Q}_{\mathbf{1}}^{\mathsf{T}} \boldsymbol{h}^i(\mathbf{X}) - \mathcal{Q}_{\mathbf{1}}^{\mathsf{T}} \boldsymbol{\rho}^i\|_{\sigma_\rho^2}^2 \tag{6.12}$$

and The third cost function

$$\|\boldsymbol{r}_3(\mathbf{X}, \mathbf{N})\|^2 \triangleq \sum_i \|\mathcal{Q}_{\mathbf{1}}^{\mathsf{T}} \boldsymbol{h}^i(\mathbf{X}) + \lambda \mathcal{R}_{\mathbf{1}} N^i - \mathcal{Q}_{\mathbf{1}}^{\mathsf{T}} \boldsymbol{\varphi}^i\|_{\sigma_\varphi^2}^2 \tag{6.13}$$

will define the common position shift and is analyzed in Section 6.2.2; therefore

$$\|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2 = \|\boldsymbol{r}_1(\mathbf{X})\|^2 + \|\boldsymbol{r}_2(\mathbf{X})\|^2 + \|\boldsymbol{r}_3(\mathbf{X}, \mathbf{N})\|^2. \tag{6.14}$$

To simplify expressions in the following discussion, let

$$\|\boldsymbol{r}_a(\mathbf{X})\|^2 \triangleq \|\boldsymbol{r}_1(\mathbf{X})\|^2 + \|\boldsymbol{r}_2(\mathbf{X})\|^2$$

$$\|\boldsymbol{r}_b(\mathbf{X})\|^2 \triangleq \|\boldsymbol{r}_2(\mathbf{X})\|^2 + \|\boldsymbol{r}_3(\mathbf{X}, \mathbf{N})\|^2.$$

6.2.4   Integer-Free Solution

Define the integer-free solution as

$$\mathbf{X}^{\circledast} = \underset{\mathbf{X} \in \mathbb{R}^{n_s K}}{arg\,min} \|\boldsymbol{r}_1(\mathbf{X})\|^2. \tag{6.15}$$

145

Proposition 6.5. *If the variable* $\mathbf{N}$ *is treated as a real vector, then for* $\check{\mathbf{X}}$ *and* $\check{\mathbf{N}}$ *as defined in eqn. (5.33)*

$$\|\boldsymbol{r}_1(\mathbf{X}^\circledast)\|^2 = \|\boldsymbol{r}(\check{\mathbf{X}}, \check{\mathbf{N}})\|^2$$

and $\mathbf{X}^\circledast = \check{\mathbf{X}}$ *where* $\boldsymbol{r}(\mathbf{X}, \mathbf{N})$ *is defined in eqn. (5.33).* $\qquad\qquad \triangle$

Proof. From eqn. (6.14),

$$\|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2 = \|\boldsymbol{r}_1(\mathbf{X})\|^2 + \sum_i \left\| \mathcal{Q}_1^{\mathsf{T}} \boldsymbol{h}^i(\mathbf{X}) + \lambda \mathcal{R}_1 N^i - \mathcal{Q}_1^{\mathsf{T}} \boldsymbol{\varphi}^i \right\|_{\sigma_\varphi^2}^2 .$$

Each term in the summation is a scalar:

$$\mathcal{Q}_1^{\mathsf{T}} \boldsymbol{h}^i(\mathbf{X}) + \lambda \mathcal{R}_1 N^i - \mathcal{Q}_1^{\mathsf{T}} \boldsymbol{\varphi}^i. \qquad\qquad (6.16)$$

When N^i is treated as a real variable, then for any $\mathbf{X}$ the value

$$\check{N}^i \;=\; \frac{\mathcal{Q}_1^{\mathsf{T}} \left(\boldsymbol{\varphi}^i - \boldsymbol{h}^i(\mathbf{X}) \right)}{\lambda \mathcal{R}_1} \qquad\qquad (6.17)$$

makes the i-th term zero. Therefore,

$$\|\boldsymbol{r}(\mathbf{X}, \check{\mathbf{N}})\| = \|\boldsymbol{r}_1(\mathbf{X})\|^2.$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

Proposition 6.5 indicates that the integer-free solution is equivalent to the float solution of eqn. (5.33). This equivalence will be utilized in Section 6.3.

6.3 Common Position Shift Method Analysis

This section presents mathematical analysis of the proposed CPS method. With the aid of several propositions proved in this section, the major proposition of the CPS method is shown in Proposition 6.12.

6.3.1 Useful Constants related to GPS

In the analysis to follow, certain GPS related facts will be used [3, 36, 54]. They are summarized in this paragraph. The standard deviation of the differential pseudorange measurement is $\sigma_\rho = 0.1 \sim 3$ meters. The standard deviation of the differential phase measurement is $\sigma_\varphi \approx 0.01\sigma_\rho$. The minimum distance from a receiver on the earth surface to a GPS satellites satisfies

$$h_k^i\big(\boldsymbol{p}_k\big) = \|\boldsymbol{p}(t_k) - \boldsymbol{p}^i(t_k)\| \geq \underline{D} \triangleq 20000km.$$

The orbital speed of GPS satellite with respect to the ECEF origin satisfies $\|V^i\| \leq \bar{V} \triangleq 4.0km/s$. With a base station in the range of a few tens of kilometers, DGPS accuracy is in the order of 1 meter (i.e., 1 σ), growing at the rate of 1 meter per $150km$ of separation [36]. Herein, it is assumed that there are always proprietary or public base stations (e.g. from CORS [60]) available to the rover within $20km$. Therefore, with the aid of the carrier phase measurements, for the float solution $\check{\mathbf{X}}$, $\|\check{\boldsymbol{p}}(t_k) - \boldsymbol{p}(t_k)\| < \Delta_f = 3m.$

6.3.2 Propositions for the CPS method

Two propositions are presented to facilitate later discussion of our CPS approach.

Proposition 6.6. *Neglecting the time variation and the high-order-terms in the linearization of the GPS measurement model, for any trajectory estimate $\hat{\mathbf{X}} \in \mathbb{R}^{n_s K}$ it is valid that*

$$\|\boldsymbol{r}_1(\hat{\mathbf{X}} \oplus \Delta\boldsymbol{p})\|^2 = \|\boldsymbol{r}_1(\hat{\mathbf{X}})\|^2. \tag{6.18}$$

Accounting for the time variation and the high-order-terms in the linearization of the GPS measurement model, for any trajectory estimate $\hat{\mathbf{X}} \in \mathbb{R}^{n_s K}$ with $\|\Delta p\| < 10 km$ it is valid that

$$\bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i(\hat{\mathbf{X}} \oplus \Delta\boldsymbol{p}) - \boldsymbol{\varphi}^i \right) = \bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i(\hat{\mathbf{X}}) - \boldsymbol{\varphi}^i + \boldsymbol{\delta}_1 \right), \tag{6.19}$$

$$\bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i(\hat{\mathbf{X}} \oplus \Delta\boldsymbol{p}) - \boldsymbol{\rho}^i \right) = \bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i(\hat{\mathbf{X}}) - \boldsymbol{\rho}^i + \boldsymbol{\delta}_1 \right), \tag{6.20}$$

where $\boldsymbol{\delta}_1 \in \mathbb{R}^K$ is a vector of perturbations caused by the common position shift. Furthermore, the magnitude of $\boldsymbol{\delta}_1$ is bounded by

$$\|\boldsymbol{\delta}_1\|_\infty \le B_1(\|\Delta\boldsymbol{p}\|, \bar{v}), \tag{6.21}$$

where the real function $B_1 : \mathbb{R}_+^2 \mapsto \mathbb{R}_+$ is defined as

$$B_1(\|\Delta\boldsymbol{p}\|, \bar{v}) \triangleq K(C_1 + C_2\bar{v})\|\Delta\boldsymbol{p}\| + \|\Delta\boldsymbol{p}\|^2/2\underline{D} \tag{6.22}$$

and K is the CRT window length, $C_1 \triangleq \frac{\bar{V}}{\underline{D}} = 2.0 \times 10^{-4}$, $C_2 \triangleq \frac{1}{\underline{D}} = 5.0 \times 10^{-8} s/m$, $\bar{v}$ is the upper bound of the rover speed over the window, and $\|\Delta\boldsymbol{p}\|$ is the magnitude of common position shift. $\triangle$

Proof. First, a common-position-shift $\Delta\boldsymbol{p}$ will not cause any variation in the prior cost and the INS cost terms, i.e.,

$$\sum_k \left\| \boldsymbol{\phi}\big(\boldsymbol{x}(t_k \oplus \Delta\boldsymbol{p}), \mathbf{U}_k\big) - \boldsymbol{x}(t_{k+1} \oplus \Delta\boldsymbol{p}) \right\|_{\mathbf{Q}_k}^2 = \sum_k \left\| \boldsymbol{\phi}\big(\boldsymbol{x}(t_k), \mathbf{U}_k\big) - \boldsymbol{x}(t_{k+1}) \right\|_{\mathbf{Q}_k}^2 .$$

So, to prove eqn. (6.18), the GPS measurements are focused on in the following. Eqn. (6.19) and (6.20) are first derived by considering the time variation and the high-order-terms in the linearization of the GPS measurement model, and then by neglecting the perturbation $\boldsymbol{\delta}_1$, eqn. (6.18) can be obtained.

With the assumption $\|\Delta\boldsymbol{p}\| < 10km$, the numerical analysis in Section 8.8.1.3 of [22] show that

$$h_k^i\big(\boldsymbol{p}_k + \Delta\boldsymbol{p}\big) - h_k^i\big(\boldsymbol{p}_k\big) \approx H_k^i \Delta\boldsymbol{p} + \frac{\|\Delta\boldsymbol{p}\|^2}{2h_k^i\big(\boldsymbol{p}_k\big)} + h.o.t.,$$

where $h_k^i\big(\boldsymbol{p}_k\big) = \|\boldsymbol{p}(t_k) - \boldsymbol{p}^i(t_k)\|$ and $H_k^i \in \mathbb{R}^{1\times 3}$ is the Jacobian matrix of h_k^i:

$$H_k^i \triangleq H^i(t_k) = \frac{\partial h_k^i}{\partial \boldsymbol{p}(t_k)} = \left[\frac{\boldsymbol{p}(t_k) - \boldsymbol{p}^i(t_k)}{\|\boldsymbol{p}(t_k) - \boldsymbol{p}^i(t_k)\|} \right].$$

Given that $h_k^i\big(\boldsymbol{p}_k\big) \geq \underline{D}$ and $\epsilon_k^i \triangleq \|\Delta\boldsymbol{p}\|^2 / \big(2h_k^i\big(\boldsymbol{p}_k\big)\big)$, it follows that

$$h_k^i\big(\boldsymbol{p}_k + \Delta\boldsymbol{p}\big) - h_k^i\big(\boldsymbol{p}_k\big) = H_k^i \Delta\boldsymbol{p} + \epsilon_k^i, \tag{6.23}$$

with $|\epsilon_k^i| < \|\Delta\boldsymbol{p}\|^2 / 2\underline{D}$.

Considered as a function of time $H^i(t) : \mathbb{R} \mapsto \mathbb{R}^3$, let

$$\bar{t} \triangleq (t_K + t_1)/2 \quad \text{and} \quad \bar{H}^i \triangleq H^i(\bar{t}), \tag{6.24}$$

then by Taylor series, it follows that

$$H^i(t) = \bar{H}^i + \frac{dH^i(t)}{dt}(t - \bar{t}) + h.o.t. \tag{6.25}$$

By defining

$$\check{H}_k^i \triangleq H_k^i - \bar{H}^i,$$

and using eqn. (6.25), $\check{H}_k^i$ can be written as

$$\check{H}_k^i = \frac{dH^i(t_k)}{dt}(t_k - \bar{t}) + h.o.t.$$

The derivative $dH^i(t)/dt$ is evaluated as

$$\frac{dH^i(t_k)}{dt} = \frac{\boldsymbol{v}(t) - \boldsymbol{v}^i(t)}{\|\boldsymbol{p}(t) - \boldsymbol{p}^i(t)\|} - \frac{(\boldsymbol{p}(t) - \boldsymbol{p}^i(t))(\boldsymbol{p}(t) - \boldsymbol{p}^i(t))^{\mathsf{T}}(\boldsymbol{v}(t) - \boldsymbol{v}^i(t))}{\|\boldsymbol{p}(t) - \boldsymbol{p}^i(t)\|^3}; \tag{6.26}$$

therefore, the variation of the Jacobian matrix is bounded by

$$\|\check{H}_k^i\| \leq \frac{2\left(\bar{v} + \bar{V}\right)}{\underline{D}}(t_k - \bar{t}),$$

since $\|\boldsymbol{v}(t)\| \leq \bar{v}$ and $\|\boldsymbol{v}^i(t)\| \leq \bar{V}$. Furthermore, with Assumption 3, $(t_K - t_1) < K$

then

$$\|\check{H}_k^i\| \leq \frac{\bar{V} + \bar{v}}{\underline{D}}(t_K - t_1) = K(C_1 + C_2\bar{v}). \tag{6.27}$$

We are now in a position to consider the effect of a shift $\Delta \boldsymbol{p}$ on the value of $\bar{\mathcal{Q}}_1^\intercal \boldsymbol{h}^i\big(\hat{\mathbf{X}} \oplus \Delta \boldsymbol{p}\big)$:

$$
\begin{aligned}
& \bar{\mathcal{Q}}_1^\intercal \boldsymbol{h}^i\big(\hat{\mathbf{X}} \oplus \Delta \boldsymbol{p}\big) \\
&= \bar{\mathcal{Q}}_1^\intercal
\begin{bmatrix}
h_1^i\big(\hat{\boldsymbol{p}}_1 + \Delta \boldsymbol{p}\big) \\
\vdots \\
h_K^i\big(\hat{\boldsymbol{p}}_K + \Delta \boldsymbol{p}\big)
\end{bmatrix} \\
&= \bar{\mathcal{Q}}_1^\intercal
\begin{bmatrix}
h_1^i\big(\hat{\boldsymbol{p}}_1\big) + H_1^i \Delta \boldsymbol{p} + \epsilon_1^i \\
\vdots \\
h_K^i\big(\hat{\boldsymbol{p}}_K\big) + H_K^i \Delta \boldsymbol{p} + \epsilon_K^i
\end{bmatrix} \\
&= \bar{\mathcal{Q}}_1^\intercal
\begin{bmatrix}
h_1^i\big(\hat{\boldsymbol{p}}_1\big) \\
\vdots \\
h_K^i\big(\hat{\boldsymbol{p}}_K\big)
\end{bmatrix}
+ \bar{\mathcal{Q}}_1^\intercal
\begin{bmatrix}
H_1^i \\
\vdots \\
H_K^i
\end{bmatrix} \Delta \boldsymbol{p} + \bar{\mathcal{Q}}_1^\intercal \boldsymbol{\epsilon}^i,
\end{aligned}
$$

where for ϵ_k^i in eqn. (6.23)

$$
\boldsymbol{\epsilon}^i \triangleq [\epsilon_1^i, \ldots, \epsilon_K^i]^\intercal \in \mathbb{R}^K. \tag{6.28}
$$

With the notation of $\bar{H}^i$ and $\check{H}^i_k$, it follows that

$$
\bar{\mathcal{Q}}^{\mathsf{T}}_1 \boldsymbol{h}^i\!\left(\hat{\mathbf{X}} \oplus \Delta \boldsymbol{p}\right)
$$

$$
= \bar{\mathcal{Q}}^{\mathsf{T}}_1 \boldsymbol{h}^i\!\left(\hat{\mathbf{X}}\right) + \bar{\mathcal{Q}}^{\mathsf{T}}_1 \begin{bmatrix} \check{H}^i_1 \\ \vdots \\ \check{H}^i_K \end{bmatrix} \Delta \boldsymbol{p} + \bar{\mathcal{Q}}^{\mathsf{T}}_1 \begin{bmatrix} \bar{H}^i \Delta \boldsymbol{p} \\ \vdots \\ \bar{H}^i \Delta \boldsymbol{p} \end{bmatrix} + \bar{\mathcal{Q}}^{\mathsf{T}}_1 \boldsymbol{\epsilon}^i
$$

$$
= \bar{\mathcal{Q}}^{\mathsf{T}}_1 \left(\boldsymbol{h}^i\!\left(\hat{\mathbf{X}}\right) + \begin{bmatrix} \check{H}^i_1 \\ \vdots \\ \check{H}^i_K \end{bmatrix} \Delta \boldsymbol{p} + \boldsymbol{\epsilon}^i \right) + \mathbf{0},
$$

where the $\mathbf{0}$ term follows from the fact that the columns of $\bar{\mathcal{Q}}_1$ span the left null space of $\mathbf{1}$. Thus, it follows that

$$
\bar{\mathcal{Q}}^{\mathsf{T}}_1 \boldsymbol{h}^i\!\left(\hat{\mathbf{X}} \oplus \Delta \boldsymbol{p}\right) = \bar{\mathcal{Q}}^{\mathsf{T}}_1 \left[\boldsymbol{h}^i\!\left(\hat{\mathbf{X}}\right) + \boldsymbol{\delta}_1 \right] \tag{6.29}
$$

with the perturbation

$$
\boldsymbol{\delta}_1 \triangleq \left[\check{H}^i_1 \Delta \boldsymbol{p} + \epsilon^i_1, \ldots, \check{H}^i_K \Delta \boldsymbol{p} + \epsilon^i_K \right]^{\mathsf{T}}. \tag{6.30}
$$

Then, from eqn. (6.27) it follows that

$$
\|\boldsymbol{\delta}_1\|_\infty = \max\{|\check{H}^i_k \Delta \boldsymbol{p} + \epsilon^i_1|\}
$$

$$
\leq K(C_1 + C_2 \bar{v})\|\Delta \boldsymbol{p}\| + \|\Delta \boldsymbol{p}\|^2/2\underline{D},
$$

where $C_1 = 2.0 \times 10^{-4}$ and $C_2 = 5.0 \times 10^{-8} s/m$.

With eqn. (6.29), eqn. (6.19) and (6.20) can be derived, and then by neglecting the perturbation $\boldsymbol{\delta}_1$, eqn. (6.18) is valid. $\qquad \square$

Remark 6.7. *The proof of Proposition 6.6 implies that the linear transformation introduced by $\bar{Q}^{\intercal}$ will remove the mean from a vector, leaving the variations. In practice, the curvature of the GPS satellite orbit is relative small compared with the distance from the satellite vehicle to the earth surface. Thus, h_k^i and H^i have weak nonlinearity and vary slowly with respect to time, and then can be considered as 'locally' time-invariant and linear. Thus, the perturbation $\boldsymbol{\delta}_1$ caused by the common position shift is very small. For example, when $K \leq 10$, $\|\Delta \boldsymbol{p}\| \leq 1.5m$, $\bar{v} \leq 50m/s$ the upper bound for the perturbation $\|\check{H}_k^i \Delta \boldsymbol{p}\|$ caused by the common position shift is upper-bounded by 0.0031 meter. In this case, the perturbation magnitude is even smaller than the centimeter noise level of the carrier phase measurement.* $\triangle$

The following proposition is with respect to cost functions $\|\boldsymbol{r}_2(\mathbf{X})\|^2$ and $\|\boldsymbol{r}_3(\mathbf{X}, \mathbf{N})\|^2$ defined in eqn. (6.13).

Proposition 6.8. *Neglecting the time variation and the high-order-terms in the linearization of the GPS measurement model, for any two trajectory-integer estimates $(\hat{\mathbf{X}}_1, \hat{\mathbf{N}}_1)$ and $(\hat{\mathbf{X}}_2, \hat{\mathbf{N}}_2)$ if we have*

$$\|\boldsymbol{r}_2(\hat{\mathbf{X}}_1)\|^2 + \|\boldsymbol{r}_3(\hat{\mathbf{X}}_1, \hat{\mathbf{N}}_1)\|^2 > \|\boldsymbol{r}_2(\hat{\mathbf{X}}_2)\|^2 + \|\boldsymbol{r}_3(\hat{\mathbf{X}}_2, \hat{\mathbf{N}}_2)\|^2,$$

then there must exist a correction $(\Delta \boldsymbol{p}, \delta \mathbf{N}) \in \mathbb{R}^3 \bigcup \mathbb{Z}^m$ such that

$$\|\boldsymbol{r}_2(\hat{\mathbf{X}}_1 \oplus \Delta \boldsymbol{p})\|^2 + \|\boldsymbol{r}_3(\hat{\mathbf{X}}_1 \oplus \Delta \boldsymbol{p}, \hat{\mathbf{N}}_1 + \delta \mathbf{N})\|^2 = \|\boldsymbol{r}_2(\hat{\mathbf{X}}_2)\|^2 + \|\boldsymbol{r}_3(\hat{\mathbf{X}}_2, \hat{\mathbf{N}}_2)\|^2. \quad (6.31)$$

Define the position errors between these two trajectory estimates as $\delta \mathbf{p}_k \triangleq \boldsymbol{p}_k^2 - \boldsymbol{p}_k^1, k = 1, \ldots, K$. Accounting for the time variation and the high-order-terms in the

linearization of the GPS measurement model, for any $(\hat{\mathbf{X}}_1, \hat{\mathbf{N}}_1)$ *and* $(\hat{\mathbf{X}}_2, \hat{\mathbf{N}}_2)$ *with*

$\|\delta\boldsymbol{p}_k\| < 10km$, *there exists a correction* $(\Delta\boldsymbol{p}, \delta\mathbf{N}) \in \mathbb{R}^3 \bigcup \mathbb{Z}^m$ *such that*

$$\mathcal{Q}_1^\mathsf{T}\boldsymbol{h}^i(\hat{\mathbf{X}}_1 \oplus \Delta\boldsymbol{p}) + \lambda\mathcal{R}_1(\hat{N}_1^i + \delta N^i) - \mathcal{Q}_1^\mathsf{T}\boldsymbol{\varphi}^i$$
$$= \mathcal{Q}_1^\mathsf{T}\boldsymbol{h}^i(\hat{\mathbf{X}}_2) + \lambda\mathcal{R}_1\hat{N}_2^i - \mathcal{Q}_1^\mathsf{T}\boldsymbol{\varphi}^i + \mathcal{Q}_1^\mathsf{T}\boldsymbol{\delta}_2, \quad (6.32)$$

$$\mathcal{Q}_1^\mathsf{T}\boldsymbol{h}^i(\hat{\mathbf{X}}_1 \oplus \Delta\boldsymbol{p}) - \mathcal{Q}_1^\mathsf{T}\boldsymbol{\rho}^i = \mathcal{Q}_1^\mathsf{T}\boldsymbol{h}^i(\mathbf{X}_2) - \mathcal{Q}_1^\mathsf{T}\boldsymbol{\rho}^i + \mathcal{Q}_1^\mathsf{T}\boldsymbol{\delta}_2, \quad (6.33)$$

where the magnitude of $\boldsymbol{\delta}_2$ *is bounded by*

$$\|\boldsymbol{\delta}_2\|_\infty \le B_1(\|\Delta\boldsymbol{p}\|, \bar{v}), \quad (6.34)$$

B_1 *is defined in eqn. (6.22),* $\bar{v}$ *is the upper bound of the rover speed over the window,*

and $\|\Delta\boldsymbol{p}\|$ *is the magnitude of common position shift.* $\triangle$

Proof. For for the sake of saving space, the linearization error term $\boldsymbol{\epsilon}^i$ is not included in this proof, but can be added by following the idea in the proof of Proposition 6.6.

To show the existence of $(\Delta\boldsymbol{p}, \mathbf{N})$ for eqn. (6.31), we can first show the existence of $(\Delta\boldsymbol{p}, \mathbf{N})$ for eqn. (6.32) and (6.33), then neglect $\boldsymbol{\delta}_2$.

Based on our construction in Section 6.2, $\mathcal{Q}_1$ is a column matrix and its column is a normal vector spans the range space of $\mathbf{1} = [1, \dots, 1]^\mathsf{T} \in \mathbb{R}^K$, i.e.

$$\mathcal{Q}_1^\mathsf{T} = [1/\sqrt{K}, \dots, 1/\sqrt{K}] \in \mathbb{R}^K.$$

Based on the definition of $\mathcal{Q}_1$, we have that

$$\mathcal{Q}_1^\mathsf{T} \begin{bmatrix} \bar{H}^i \\ \vdots \\ \bar{H}^i \end{bmatrix} = \sqrt{K}\bar{H}^i \tag{6.35}$$

for $\bar{H}^i$ is defined in eqn. (6.24). This experession will be used in the following.

Starting from eqn. (6.32) our goal is to show that

$$\mathcal{Q}_1^\mathsf{T} \boldsymbol{h}^i\big(\hat{\mathbf{X}}_1 \oplus \Delta \boldsymbol{p}\big) + \lambda \mathcal{R}_1(\hat{N}_1^i + \delta N^i) = \mathcal{Q}_1^\mathsf{T} \boldsymbol{h}^i\big(\hat{\mathbf{X}}_2\big) + \lambda \mathcal{R}_1 N_2^i + \mathcal{Q}_1^\mathsf{T} \boldsymbol{\delta}_2. \tag{6.36}$$

Let $\delta\mathbf{X} \triangleq \hat{\mathbf{X}}_2 - \hat{\mathbf{X}}_1$ and $\hat{\boldsymbol{p}}_k \in \mathbb{R}^3$ be the position estimates in $\hat{\mathbf{X}}_1$, then

$$\mathcal{Q}_1^\mathsf{T} \boldsymbol{h}^i\big(\hat{\mathbf{X}}_2\big) \tag{6.37}$$

$$= \mathcal{Q}_1^\mathsf{T} \boldsymbol{h}^i\big(\hat{\mathbf{X}}_1 + \delta\mathbf{X}\big) \tag{6.38}$$

$$= \mathcal{Q}_1^\mathsf{T} \begin{bmatrix} h_1^i\big(\hat{\boldsymbol{p}}_1 + \Delta\boldsymbol{p}_1\big) \\ \vdots \\ h_K^i\big(\hat{\boldsymbol{p}}_K + \Delta\boldsymbol{p}_K\big) \end{bmatrix} \tag{6.39}$$

$$= \mathcal{Q}_1^\mathsf{T} \begin{bmatrix} h_1^i\big(\hat{\boldsymbol{p}}_1\big) \\ \vdots \\ h_K^i\big(\hat{\boldsymbol{p}}_K\big) \end{bmatrix} + \mathcal{Q}_1^\mathsf{T} \begin{bmatrix} H_1^i \Delta\boldsymbol{p}_1 \\ \vdots \\ H_K^i \Delta\boldsymbol{p}_K \end{bmatrix} \tag{6.40}$$

$$= \mathcal{Q}_1^\mathsf{T} \begin{bmatrix} h_1^i\big(\hat{\boldsymbol{p}}_1\big) \\ \vdots \\ h_K^i\big(\hat{\boldsymbol{p}}_K\big) \end{bmatrix} + \mathcal{Q}_1^\mathsf{T} \begin{bmatrix} \bar{H}^i \Delta\boldsymbol{p}_1 \\ \vdots \\ \bar{H}^i \Delta\boldsymbol{p}_K \end{bmatrix} + \mathcal{Q}_1^\mathsf{T} \begin{bmatrix} \check{H}_1^i \Delta\boldsymbol{p}_1 \\ \vdots \\ \check{H}_K^i \Delta\boldsymbol{p}_K \end{bmatrix} \tag{6.41}$$

$$= \mathcal{Q}_1^\mathsf{T} \begin{bmatrix} h_1^i\big(\hat{\boldsymbol{p}}_1\big) \\ \vdots \\ h_K^i\big(\hat{\boldsymbol{p}}_K\big) \end{bmatrix} + \frac{\bar{H}^i}{\sqrt{K}} \sum_k \Delta\boldsymbol{p}_k + \mathcal{Q}_1^\mathsf{T} \begin{bmatrix} \check{H}_1^i(\Delta\boldsymbol{p} + \Delta\check{\boldsymbol{p}}_1) \\ \vdots \\ \check{H}_K^i(\Delta\boldsymbol{p} + \Delta\check{\boldsymbol{p}}_K) \end{bmatrix}$$

where $\bar{H}^i$ is the average of Jacobian matrix H_k^i,

$$\Delta\boldsymbol{p} = \frac{1}{K} \sum_{i=1}^{K} \Delta\boldsymbol{p}_k \in \mathbb{R}^3 \tag{6.42}$$

156

is the average of the position adjustments in $\delta\mathbf{X}$, and $\Delta\check{\boldsymbol{p}}_k$ is the variation of the 3D position adjustments. Continuing, using eqn. (6.35):

$$\mathcal{Q}_1^{\mathsf{T}}\boldsymbol{h}^i(\hat{\mathbf{X}}_2) \tag{6.43}$$

$$= \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} h_1^i(\hat{\boldsymbol{p}}_1) \\ \vdots \\ h_K^i(\hat{\boldsymbol{p}}_K) \end{bmatrix} + \sqrt{K}\bar{H}^i\Delta\boldsymbol{p} + \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} \check{H}_1^i(\Delta\boldsymbol{p}+\Delta\check{\boldsymbol{p}}_1) \\ \vdots \\ \check{H}_K^i(\Delta\boldsymbol{p}+\Delta\check{\boldsymbol{p}}_K) \end{bmatrix}$$

$$= \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} h_1^i(\hat{\boldsymbol{p}}_1) \\ \vdots \\ h_K^i(\hat{\boldsymbol{p}}_K) \end{bmatrix} + \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} (\bar{H}^i+\check{H}_1^i)\Delta\boldsymbol{p} \\ \vdots \\ (\bar{H}^i+\check{H}_K^i)\Delta\boldsymbol{p} \end{bmatrix} + \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} \check{H}_1^i\Delta\check{\boldsymbol{p}}_1 \\ \vdots \\ \check{H}_K^i\Delta\check{\boldsymbol{p}}_K \end{bmatrix}$$

$$= \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} h_1^i(\hat{\boldsymbol{p}}_1)+H_1^i\Delta\boldsymbol{p} \\ \vdots \\ h_K^i(\hat{\boldsymbol{p}}_K)+H_K^i\Delta\boldsymbol{p} \end{bmatrix} + \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} \check{H}_1^i\Delta\check{\boldsymbol{p}}_1 \\ \vdots \\ \check{H}_K^i\Delta\check{\boldsymbol{p}}_K \end{bmatrix} \tag{6.44}$$

$$= \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} h_1^i(\hat{\boldsymbol{p}}_1+\Delta\boldsymbol{p}) \\ \vdots \\ h_K^i(\hat{\boldsymbol{p}}_K+\Delta\boldsymbol{p}) \end{bmatrix} + \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} \check{H}_1^i\Delta\check{\boldsymbol{p}}_1 \\ \vdots \\ \check{H}_K^i\Delta\check{\boldsymbol{p}}_K \end{bmatrix} \tag{6.45}$$

$$= \mathcal{Q}_1^{\mathsf{T}}\boldsymbol{h}^i(\hat{\mathbf{X}}_1\oplus\Delta\boldsymbol{p}) + \mathcal{Q}_1^{\mathsf{T}}\begin{bmatrix} \check{H}_1^i\Delta\check{\boldsymbol{p}}_1 \\ \vdots \\ \check{H}_K^i\Delta\check{\boldsymbol{p}}_K \end{bmatrix}. \tag{6.46}$$

Thus, with $\hat{N}_1^i + \delta N^i = \hat{N}_2^i$, it follows that

$$\mathcal{Q}_1^{\mathsf{T}}\boldsymbol{h}^i(\hat{\mathbf{X}}_1\oplus\Delta\boldsymbol{p}) + \lambda\mathcal{R}_1(\hat{N}_1^i+\delta N^i) = \mathcal{Q}_1^{\mathsf{T}}\boldsymbol{h}^i(\hat{\mathbf{X}}_2) + \lambda\mathcal{R}_1\hat{N}_2^i + \mathcal{Q}_1^{\mathsf{T}}\boldsymbol{\delta}_2,$$

where $\boldsymbol{\delta}_2 \triangleq [\check{H}_1^i \Delta \check{\boldsymbol{p}}_1, \ldots, \check{H}_K^i \Delta \check{\boldsymbol{p}}_K]^\intercal$. With the eqn. (6.27), $\|\Delta \check{\boldsymbol{p}}_k\| < \Delta_p$ and $|\epsilon_k^i| < \|\Delta \boldsymbol{p}\|^2 / 2\underline{D}$, the bound in eqn. (6.34) can be derived. This conclusion can also apply to code measurements such that eqn. (6.33) is also valid.

By neglecting the perturbation $\boldsymbol{\delta}_2$ in eqn. (6.32-6.33), eqn. (6.31) can be obtained and this concludes the proof. $\qquad\square$

Remark 6.9. *Proposition 6.8 shows when K and Δ_p are bounded (e.g. $K \leq 10$, $\Delta_p \leq 3$), we can minimize $\|\boldsymbol{r}_3(\mathbf{X}, \mathbf{N})\|^2$ within a bound, just through a common position shift and adjusting the integer estimates. Furthermore, the magnitude of the error $\boldsymbol{\delta}_2$ is even smaller than the noise level of carrier phase measurements if K and Δ_p is small enough (e.g. when $K \leq 10$, $\Delta_p \leq 1.5m$ and $\bar{v} \leq 50$, it follows that $\|\boldsymbol{\delta}_2\|_\infty \leq 0.0031$).* $\qquad\triangle$

6.3.3   Optimality of the CPS method

The major proposition about the optimality of the CPS method is presented as follows.

Proposition 6.10. *If $\boldsymbol{\delta}_1 = \mathbf{0}$ and $\boldsymbol{\delta}_2 = \mathbf{0}$, then the following identity is valid,*

$$\|\boldsymbol{r}(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star, \mathbf{N}^\star)\|^2 = \|\boldsymbol{r}(\mathbf{X}^*, \mathbf{N}^*)\|^2, \qquad (6.47)$$

where $\check{\mathbf{X}}$ is the float solution from eqn. (5.33), $(\Delta \boldsymbol{p}^\star, \mathbf{N}^\star)$ is the CPS solution from eqn. (6.9) and $(\mathbf{X}^*, \mathbf{N}^*)$ is the full NMILS estimate from eqn. (6.6). $\qquad\triangle$

The proof of Proposition 6.10. will use the following symbols:

$$\mathbf{X}_1^* = \underset{\mathbf{X} \in \mathbb{R}^{n_s K}}{arg\,min} \|\boldsymbol{r}_1(\mathbf{X})\|^2, \quad \mathbf{X}_2^* = \underset{\mathbf{X} \in \mathbb{R}^{n_s K}}{arg\,min} \|\boldsymbol{r}_2(\mathbf{X})\|^2,$$

$$(\mathbf{X}_\star, \mathbf{N}_\star) \triangleq \underset{\mathbf{X} \in \mathbb{R}^{n_s K}, \mathbf{N} \in \mathbb{Z}^m}{arg\,min} \|\boldsymbol{r}_b(\mathbf{X}, \mathbf{N})\|^2,$$

and

$$\Delta\boldsymbol{p}_{12}^* = \underset{\Delta\boldsymbol{p} \in \mathbb{R}^3}{arg\,min} \|\boldsymbol{r}_2(\mathbf{X}_1^* \oplus \Delta\boldsymbol{p})\|^2.$$

Note that $\mathbf{X}_1^*$, $\mathbf{X}_2^*$ and $\mathbf{X}_\star$ are sets of trajectories. At each time t, with the definition in eqn. (6.1), one trajectory $\mathbf{X}$ can be rewritten as $\mathbf{X} = [\mathbf{P}^\mathsf{T}, \mathbf{S}^\mathsf{T}]^\mathsf{T}$, where $\mathbf{P} = [\boldsymbol{p}^\mathsf{T}(t_1), \ldots, \boldsymbol{p}^\mathsf{T}(t_K)]^\mathsf{T} \in \mathbb{R}^{3K}$ and $\mathbf{S} = [\boldsymbol{s}^\mathsf{T}(t_1), \ldots, \boldsymbol{s}^\mathsf{T}(t_K)]^\mathsf{T} \in \mathbb{R}^{K(n_s-3)}$. Because by definition $\|\boldsymbol{r}_2\|^2$ (or $\|\boldsymbol{r}_b\|^2$) is independent of $\mathbf{S}$, the set $\mathbf{X}_2^*$ (or $\mathbf{X}_\star$) contains all trajectories with the same sequence of positions $\mathbf{P}_2^*$, but distinct values of $\mathbf{S}$. Each trajectory in $\mathbf{X}_2^*$ (or $\mathbf{X}_\star$) has the same value for $\|\boldsymbol{r}_2\|^2$ (or $\|\boldsymbol{r}_b\|^2$), but will be penalized differently by $\|\boldsymbol{r}_1\|^2$. Similarly, $\mathbf{X}_1^*$ is a trajectory set, each having the same shape $\mathbf{S}_1^*$, but with $\mathbf{P}$ shifted by a common vector $\Delta\boldsymbol{p}$ by Proposition 6.6. Each trajectory in $\mathbf{X}_1^*$ has the same value for $\|\boldsymbol{r}_1\|^2$ but different penalty for $\|\boldsymbol{r}_b\|^2$.

Proof. By the definition of $(\mathbf{X}_\star, \mathbf{N}_\star)$, it follows that for the float solution $\check{\mathbf{X}}$ defined in eqn. (5.33), $\forall(\Delta\boldsymbol{p}, \mathbf{N}) \in \mathbb{R}^3 \cup \mathbb{Z}^m$,

$$\|\boldsymbol{r}_b(\mathbf{X}_\star, \mathbf{N}_\star)\|^2 \le \|\boldsymbol{r}_b(\check{\mathbf{X}} \oplus \Delta\boldsymbol{p}, \mathbf{N})\|^2. \tag{6.48}$$

Given $(\hat{\mathbf{X}}, \hat{\mathbf{N}})$ and $(\mathbf{X}_\star, \mathbf{N}_\star)$, the unique $(\Delta\boldsymbol{p}^\star, \mathbf{N}^\star)$ such that

$$\|\boldsymbol{r}_b(\check{\mathbf{X}} \oplus \Delta\boldsymbol{p}^\star, \mathbf{N}^\star)\|^2 = \|\boldsymbol{r}_b(\mathbf{X}_\star, \mathbf{N}_\star)\|^2, \tag{6.49}$$

is provided in Appendix III in the proof of Proposition 6.8 and can be found by the CPS estimation in eqn. (6.9). From Proposition 6.6 (with $\boldsymbol{\delta}_1 = \mathbf{0}$), it follows that

$$\|\boldsymbol{r}_1(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star)\|^2 = \|\boldsymbol{r}_1(\check{\mathbf{X}})\|^2. \tag{6.50}$$

Similarly, by Proposition 6.6 (with $\boldsymbol{\delta}_1 = \mathbf{0}$), it can also be proven that

$$\|\boldsymbol{r}_1(\mathbf{X}_1^* \oplus \Delta \boldsymbol{p}_{12}^*)\|^2 = \|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2. \tag{6.51}$$

By Proposition 3 (with $\boldsymbol{\delta}_2 = \mathbf{0}$) and considering the identical optimality achieved by $\boldsymbol{p}_{12}^*$ and $\mathbf{X}_2^*$, it follows

$$\|\boldsymbol{r}_2(\mathbf{X}_1^* \oplus \Delta \boldsymbol{p}_{12}^*)\|^2 = \|\boldsymbol{r}_2(\mathbf{X}_2^*)\|^2. \tag{6.52}$$

Since $\check{\mathbf{X}}$ is the float solution such that

$$\|\boldsymbol{r}_a(\check{\mathbf{X}})\|^2 \leq \|\boldsymbol{r}_a(\mathbf{X})\|^2, \quad \forall \mathbf{X} \in \mathbb{R}^{n_s K},$$

we have

$$\|\boldsymbol{r}_1(\check{\mathbf{X}})\|^2 + \|\boldsymbol{r}_2(\check{\mathbf{X}})\|^2 \leq \|\boldsymbol{r}_1(\mathbf{X}_1^* \oplus \Delta \boldsymbol{p}_{12}^*)\|^2 + \|\boldsymbol{r}_2(\mathbf{X}_1^* \oplus \Delta \boldsymbol{p}_{12}^*)\|^2. \tag{6.53}$$

Substituting eqn. (6.51-6.52) into (6.53), it follows that

$$\|\boldsymbol{r}_1(\check{\mathbf{X}})\|^2 + \|\boldsymbol{r}_2(\check{\mathbf{X}})\|^2 \leq \|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2 + \|\boldsymbol{r}_2(\mathbf{X}_2^*)\|^2.$$

By the definition of $\mathbf{X}_1^*$ and $\mathbf{X}_2^*$,

$$\|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2 \leq \|\boldsymbol{r}_1(\mathbf{X})\|^2, \quad \forall \mathbf{X} \in \mathbb{R}^{n_s K},$$

$$\|\boldsymbol{r}_2(\mathbf{X}_2^*)\|^2 \le \|\boldsymbol{r}_2(\mathbf{X})\|^2, \quad \forall \mathbf{X} \in \mathbb{R}^{n_s K}.$$

Combining these three inequalities yields

$$\|\boldsymbol{r}_1(\check{\mathbf{X}})\|^2 = \|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2. \tag{6.54}$$

Since $\|\boldsymbol{r}\|^2 = \|\boldsymbol{r}_1\|^2 + \|\boldsymbol{r}_b\|^2$, by the definition of $\mathbf{X}_1^*$ and $(\mathbf{X}_\star, \mathbf{N}_\star)$ we have

$$\|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2 + \|\boldsymbol{r}_b(\mathbf{X}_\star, \mathbf{N}_\star)\|^2 \le \|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2, \ \forall(\mathbf{X}, \mathbf{N}).$$

Then, it follows that

$$\|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2 + \|\boldsymbol{r}_b(\mathbf{X}_\star, \mathbf{N}_\star)\|^2 \le \|\boldsymbol{r}(\mathbf{X}^*, \mathbf{N}^*)\|^2. \tag{6.55}$$

From eqn. (6.49), (6.50) and (6.54), we have

$$\|\boldsymbol{r}_1(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star)\|^2 + \|\boldsymbol{r}_b(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star, \mathbf{N}^\star)\|^2 \le \|\boldsymbol{r}(\mathbf{X}^*, \mathbf{N}^*)\|^2, \tag{6.56}$$

i.e.,

$$\|\boldsymbol{r}(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star, \mathbf{N}^\star)\|^2 \le \|\boldsymbol{r}(\mathbf{X}^*, \mathbf{N}^*)\|^2. \tag{6.57}$$

On the other hand, from eqn. (6.6) we have

$$\|\boldsymbol{r}(\mathbf{X}^*, \mathbf{N}^*)\|^2 \le \|\boldsymbol{r}(\mathbf{X}, \mathbf{N})\|^2, \quad \forall \mathbf{X} \in \mathbb{R}^{n_s K}, \mathbf{N} \in \mathbb{Z}^m.$$

Thus, only equality can stand

$$\|\boldsymbol{r}(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star, \mathbf{N}^\star)\|^2 = \|\boldsymbol{r}(\mathbf{X}^*, \mathbf{N}^*)\|^2,$$

and this concludes the proof. □

Remark 6.11. *This paper and proof introduce different trajectories $\check{\mathbf{X}}$, $\mathbf{X}^*$, $\mathbf{X}^{\circledast}$ and trajectory sets $\mathbf{X}_1^*$, $\mathbf{X}_2^*$ and $\mathbf{X}_\star$. The proof shows that certain components of the cost function have the same value when evaluated for different trajectories of trajectory sets Taking advantage of this allows definition of the CPS algorithm described in Section 6.9 that vastly reduces the computational load as summarized in Table I.* △

Proposition 6.10 considers the case where the linearization errors do not exist. Proposition 6.12 analyzes the effect of the linearization errors.

Proposition 6.12. *Accounting for the time variation and the high-order-terms in the linearization of the GPS measurement model, the following inequality is valid*

$$\mathbf{E}\{\|\mathbf{r}(\check{\mathbf{X}}\oplus\Delta\mathbf{p}^\star,\mathbf{N}^\star)\|^2\} \leq (1+C_3)\mathbf{E}\{\|\mathbf{r}(\mathbf{X}^*,\mathbf{N}^*)\|^2\},$$

where $\check{\mathbf{X}}$ is the float solution from eqn. (5.33), $(\Delta\mathbf{p}^\star,\mathbf{N}^\star)$ is the CPS solution from eqn. (6.9) with $\|\Delta\mathbf{p}^\star\| \leq \Delta_f$, and

$$C_3 = \frac{Km\left(4\sigma_\rho^{-2}+3\sigma_\varphi^{-2}\right)[B_1(\Delta_f,\bar{v})]^2}{(2K-1)m-3},$$

$(\mathbf{X}^*,\mathbf{N}^*)$ is the full NMILS estimate from eqn. (6.6) and $\mathbf{E}\{\cdot\}$ is the expectation operator. △

This appendix proves Proposition 6.12 using similar techniques as were used for proving Proposition 6.10. In this proof, $\check{\mathbf{X}}$ and $\Delta\mathbf{p}^*$ are known and fixed.

Proof. Proposition 6.5 shows that the float solution $\check{\mathbf{X}}$ is equal to the integer-free solution, which optimizes the cost $\|\boldsymbol{r}_a(\mathbf{X})\|^2$ such that

$$\|\boldsymbol{r}_a(\check{\mathbf{X}})\|^2 \leq \|\boldsymbol{r}_a(\mathbf{X})\|^2, \quad \forall \mathbf{X} \in \mathbb{R}^{n_s K}.$$

Taking the norm of eqn. (6.19) after replacing $\boldsymbol{\delta}_1$ with (6.30), it follows that

$$
\begin{aligned}
&\| \bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i \big(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star\big) - \boldsymbol{\varphi}^i \right) \|_{\sigma_\varphi^2 \mathbf{I}^-}^2 \\
&= \| \bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i \big(\check{\mathbf{X}}\big) - \boldsymbol{\varphi}^i \right) \|_{\sigma_\varphi^2 \mathbf{I}^-}^2 \\
&\quad + 2 \begin{bmatrix} \check{H}_1^i \Delta \boldsymbol{p}^\star + \epsilon_1^i \\ \vdots \\ \check{H}_K^i \Delta \boldsymbol{p}^\star + \epsilon_K^i \end{bmatrix}^\intercal \bar{\mathcal{Q}}_1 (\sigma_\varphi^2 \mathbf{I}^-)^{-1} \bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i \big(\check{\mathbf{X}}\big) - \boldsymbol{\varphi}^i \right) + \left\| \bar{\mathcal{Q}}_1^\intercal \begin{bmatrix} \check{H}_1^i \Delta \boldsymbol{p}^\star + \epsilon_1^i \\ \vdots \\ \check{H}_K^i \Delta \boldsymbol{p}^\star + \epsilon_K^i \end{bmatrix} \right\|_{\sigma_\varphi^2 \mathbf{I}^-}^2
\end{aligned}
\tag{6.58}
$$

Note that $\boldsymbol{E}\{\bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i \big(\check{\mathbf{X}}\big) - \boldsymbol{\varphi}^i \right)\} = \mathbf{0}$ since $\check{\mathbf{X}}$ is the float solution. Thus, we have

$$
\begin{aligned}
\boldsymbol{E} \left\{ \| \bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i \big(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star\big) - \boldsymbol{\varphi}^i \right) \|_{\sigma_\varphi^2 \mathbf{I}^-}^2 \right\} \leq \\
\boldsymbol{E} \left\{ \| \bar{\mathcal{Q}}_1^\intercal \left(\boldsymbol{h}^i \big(\check{\mathbf{X}}\big) - \boldsymbol{\varphi}^i \right) \|_{\sigma_\varphi^2 \mathbf{I}^-}^2 \right\} + \sigma_\varphi^{-2} K [B_1(\|\Delta \boldsymbol{p}^\star\|, \bar{v})]^2,
\end{aligned}
\tag{6.59}
$$

since

$$
\left\| \bar{\mathcal{Q}}_1^\intercal \begin{bmatrix} \check{H}_1^i \Delta \boldsymbol{p}^\star + \epsilon_1^i \\ \vdots \\ \check{H}_K^i \Delta \boldsymbol{p}^\star + \epsilon_K^i \end{bmatrix} \right\|_{\sigma_\varphi^2 \mathbf{I}^-}^2 \leq \left\| \mathcal{Q}^\intercal \begin{bmatrix} \check{H}_1^i \Delta \boldsymbol{p}^\star + \epsilon_1^i \\ \vdots \\ \check{H}_K^i \Delta \boldsymbol{p}^\star + \epsilon_K^i \end{bmatrix} \right\|_{\sigma_\varphi^2 \mathbf{I}}^2
$$

and $|\check{H}_k \Delta \boldsymbol{p}^\star + \epsilon_k^i| \leq B_1(\Delta \boldsymbol{p}^\star, \bar{v})$, for any $k = 1, \ldots, K$. Note that an inequality analogous to ineq. (6.59) also applies to pseudorange measurements, i.e.,

$$\boldsymbol{E}\left\{\|\bar{\mathcal{Q}}_1^\mathsf{T}\left(\boldsymbol{h}^i(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star) - \boldsymbol{\rho}^i\right)\|_{\sigma_\rho^2 \mathbf{I}-}^2\right\} \leq$$
$$\boldsymbol{E}\left\{\|\bar{\mathcal{Q}}_1^\mathsf{T}\left(\boldsymbol{h}^i(\check{\mathbf{X}}) - \boldsymbol{\rho}^i\right)\|_{\sigma_\rho^2 \mathbf{I}-}^2\right\} + \sigma_\rho^{-2} K[B_1(\|\Delta \boldsymbol{p}^\star\|, \bar{v})]^2. \quad (6.60)$$

Since a common-position-shift $\Delta \boldsymbol{p}$ will not cause any variation in the prior cost and the INS cost terms, it follows that

$$\boldsymbol{E}\left\{\|\boldsymbol{r}_1(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star)\|^2 - \|\boldsymbol{r}_1(\check{\mathbf{X}})\|^2\right\} \leq m\left(\sigma_\rho^{-2} + \sigma_\varphi^{-2}\right) K[B_1(\|\Delta \boldsymbol{p}^\star\|, \bar{v})]^2, \quad (6.61)$$

where m is the number of available satellites.

Similarly, with Proposition 6.8 and its proof, it can be derived that

$$\boldsymbol{E}\left\{\|\boldsymbol{r}_b(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star, \mathbf{N}^\star)\|^2 - \|\boldsymbol{r}_b(\mathbf{X}_\star, \mathbf{N}_\star)\|^2\right\} \leq m\left(\sigma_\rho^{-2} + \sigma_\varphi^{-2}\right) K[B_1(\|\Delta \boldsymbol{p}^\star\|, \bar{v})]^2.$$
$$(6.62)$$

Similarly by Proposition 6.6, it can also be proven that

$$\boldsymbol{E}\left\{\|\boldsymbol{r}_1(\mathbf{X}_1^* \oplus \Delta \boldsymbol{p}_{12}^*)\|^2 - \|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2\right\} \leq m\left(\sigma_\rho^{-2} + \sigma_\varphi^{-2}\right) K[B_1(\|\Delta \boldsymbol{p}_{12}^*\|, \bar{v})]^2. \quad (6.63)$$

By Proposition 6.8,

$$\boldsymbol{E}\left\{\|\boldsymbol{r}_2(\mathbf{X}_1^* \oplus \Delta \boldsymbol{p}_{12}^*)\|^2 - \|\boldsymbol{r}_2(\mathbf{X}_2^*)\|^2\right\} \leq m\sigma_\rho^{-2} K[B_1(\|\Delta \boldsymbol{p}_{12}^*\|, \bar{v})]^2. \quad (6.64)$$

Since $\check{\mathbf{X}}$ is the float solution such that

$$\|\boldsymbol{r}_a(\check{\mathbf{X}})\|^2 \leq \|\boldsymbol{r}_a(\mathbf{X})\|^2, \quad \forall \mathbf{X} \in \mathbb{R}^{n_s K},$$

we have

$$\|\boldsymbol{r}_1(\check{\mathbf{X}})\|^2 + \|\boldsymbol{r}_2(\check{\mathbf{X}})\|^2 \leq \|\boldsymbol{r}_1(\mathbf{X}_1^* \oplus \Delta\boldsymbol{p}_{12}^*)\|^2 + \|\boldsymbol{r}_2(\mathbf{X}_1^* \oplus \Delta\boldsymbol{p}_{12}^*)\|^2. \tag{6.65}$$

With ineq. (6.51-6.52) and (6.65), it follows that

$$\boldsymbol{E}\left\{\|\boldsymbol{r}_1(\check{\mathbf{X}})\|^2 + \|\boldsymbol{r}_2(\check{\mathbf{X}})\|^2 - \|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2 - \|\boldsymbol{r}_2(\mathbf{X}_2^*)\|^2\right\}$$

$$\leq m\left(2\sigma_\rho^{-2} + \sigma_\varphi^{-2}\right)K[B_1(\|\Delta\boldsymbol{p}_{12}^*\|, \bar{v})]^2.$$

and then

$$\boldsymbol{E}\left\{\|\boldsymbol{r}_1(\check{\mathbf{X}})\|^2 - \|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2\right\} \leq m\left(2\sigma_\rho^{-2} + \sigma_\varphi^{-2}\right)K[B_1(\|\Delta\boldsymbol{p}_{12}^*\|, \bar{v})]^2, \tag{6.66}$$

since $\|\boldsymbol{r}_2(\check{\mathbf{X}})\|^2 - \|\boldsymbol{r}_2(\mathbf{X}_2^*)\|^2 \geq 0$ is always true. From ineq. (6.61) and (6.66), it follows that

$$\boldsymbol{E}\left\{\|\boldsymbol{r}_1(\check{\mathbf{X}} \oplus \Delta\boldsymbol{p}^\star)\|^2 - \|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2\right\} \leq m\left(3\sigma_\rho^{-2} + 2\sigma_\varphi^{-2}\right)K[B_1(\Delta, \bar{v})]^2,$$

where $\Delta \triangleq \max(\|\Delta\boldsymbol{p}^\star\|, \|\Delta\boldsymbol{p}_{12}^*\|)$. Along with the ineq. (6.62) and the inequality

$$\|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2 + \|\boldsymbol{r}_b(\mathbf{X}_\star, \mathbf{N}_\star)\|^2 \leq \|\boldsymbol{r}(\mathbf{X}^*, \mathbf{N}^*)\|^2,$$

it follows that

$$\boldsymbol{E}\{\|\boldsymbol{r}(\check{\mathbf{X}} \oplus \Delta\boldsymbol{p}^\star, \mathbf{N}^\star)\|^2\}$$

$$= \boldsymbol{E}\left\{\|\boldsymbol{r}_1(\check{\mathbf{X}} \oplus \Delta\boldsymbol{p}^\star)\|^2 + \|\boldsymbol{r}_b(\check{\mathbf{X}} \oplus \Delta\boldsymbol{p}^\star, \mathbf{N}^\star)\|^2\right\}$$

$$\leq \boldsymbol{E}\left\{\|\boldsymbol{r}_1(\mathbf{X}_1^*)\|^2 + \|\boldsymbol{r}_b(\mathbf{X}_\star, \mathbf{N}_\star)\|^2\right\} + m\left(4\sigma_\rho^{-2} + 3\sigma_\varphi^{-2}\right)K[B_1(\Delta, \bar{v})]^2. \tag{6.67}$$

165

Following the steps from eqn. (6.55-6.57), it yields that

$$\boldsymbol{E}\{\|\boldsymbol{r}(\check{\mathbf{X}}\oplus\Delta\boldsymbol{p}^\star,\mathbf{N}^\star)\|^2\}$$

$$\leq \boldsymbol{E}\{\|\boldsymbol{r}(\mathbf{X}^*,\mathbf{N}^*)\|^2\} + m\left(4\sigma_\rho^{-2} + 3\sigma_\varphi^{-2}\right)K[B_1(\Delta,\bar{v})]^2.$$

$$= \boldsymbol{E}\{\|\boldsymbol{r}(\mathbf{X}^*,\mathbf{N}^*)\|^2\}\left(1 + C_3\right), \qquad\qquad (6.68)$$

where

$$C_3 = \frac{m\left(4\sigma_\rho^{-2} + 3\sigma_\varphi^{-2}\right)K[B_1(\Delta,\bar{v})]^2}{\boldsymbol{E}\{\|\boldsymbol{r}(\mathbf{X}^*,\mathbf{N}^*)\|^2\}}$$

The optimum $\|\boldsymbol{r}(\mathbf{X}^*,\mathbf{N}^*)\|^2$ is referred as the *a-posteriori* variance factor in classical least square literature [52]. It can be shown that

$$\boldsymbol{E}\{\|\boldsymbol{r}(\mathbf{X}^*,\mathbf{N}^*)\|^2\} = (2K-1)m - 3,$$

where $(2K-1)m - 3 = (n_s K + 2Km - 3) - (n_s K + m)$ is the difference between the total number of measurements and the total dimension of unknown variables, which is also referred as the *degree-of-freedom*. Thus, it follows that

$$C_3 = \frac{m\left(4\sigma_\rho^{-2} + 3\sigma_\varphi^{-2}\right)K[B_1(\Delta,\bar{v})]^2}{(2K-1)m - 3}$$

Herein, both $\|\Delta\boldsymbol{p}^\star\|$ and $\|\Delta\boldsymbol{p}_{12}^*\|$ are upper bounded by Δ_f, so $\Delta \leq \Delta_f$ and this concludes the proof. □

Remark 6.13. *When $K = 10$, $\Delta_f = 1.5$, $m = 7$, $\sigma_\rho = 1.0$ meters and $\sigma_\varphi = 0.020$ meters, $C_3 = 0.04$. Thus, Proposition 6.12 shows that the Common Position Shift*

is a valid approximation to the original full NMILS approach. The error between the expected final costs of two optimizations is bounded within 4% of the expected optimum from the full NMILS approach. Implementation results presented in Section 6.4 shows that the differences between two estimation results are unsubstantial. If the small perturbations $\boldsymbol{\delta}$ are omitted in Proposition 6.6 and 6.8 for practical implementations, it can be assumed that $\|\boldsymbol{r}_1(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star)\|^2 + \|\boldsymbol{r}_3(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star, \mathbf{N}^\star)\|^2 \approx \|\boldsymbol{r}(\mathbf{X}^*, \mathbf{N}^*)\|^2$, by following similar procedures in the proof of Proposition 6.12. This shows the intuition behind the proposed Common-Position-Shift approach. △

6.3.4 Computation Analysis of the CPS method

Table I compares the computation costs of the Direct MILS (see Section 5.4) and CPS MILS (see Section 6.2.1). In Table I, $\mathfrak{I}_1$ represents the number of (linearized) nonlinear least squares iterations are required to find in the float solution is Step 1. Similarly, $\mathfrak{I}_2$ represents the number of linearized iterations required for integer ambiguity resolution in the ILS problem of Step 2. The IMU sampling rate (e.g. $200Hz$) is f.

Compared with solving the full NMILS in eqn. (6.6) directly, the computational cost in eqn. (6.9) is significantly reduced due to the much smaller dimension of the real unknown variable $\Delta \boldsymbol{p} \in \mathbb{R}^3$ versus $\mathbf{X} \in \mathbb{R}^{n_s K}$. In particular, this dimension reduction facilitates the QR-decomposition in eqn. (5.36), see Row (2b) in Table I.

Table 6.1: Computation Comparison

Step	Process	Direct MILS	CPS MILS
1)	Float solution	$\mathcal{O}((n_s K)^3) \times \mathfrak{I}_1$	$\mathcal{O}((n_s K)^3) \times \mathfrak{I}_1$
2a)	Integrate INS	$f K n_s \times \mathfrak{I}_2$	0
2b)	QR of A	$2M(N_s)^2 \times \mathfrak{I}_2$	$2(2m)(3)^2 \times \mathfrak{I}_2$
2c)	QRZ of $\bar{\mathbf{Q}}_A^\intercal B$	$2(2Km)m^2 \times \mathfrak{I}_2$	$2(2m)m^2 \times \mathfrak{I}_2$
2d)	Integer Search	$(*) \times \mathfrak{I}_2$	$(*) \times \mathfrak{I}_2$
3)	Integer Valid.	Km	Km

The dimensions of the corresponding A matrices in the full MILS and CPS MILS are $M \times N_s$ versus $2m \times 3$, where $M \triangleq n_s K + 2mK - 3$ is the dimension of the residual vector, $N_s \triangleq n_s K$ is the total state dimension of $\mathbf{X}$, and $2m$ is the total number of GPS measurements at a single epoch (code and carrier phase). Furthermore, while each NMILS iteration of the direct approach requires the expensive INS reintegration, the CPS NMILS of (6.9) does not, see Row (2a) in Table I. In the ILS, the reduction step is cast as a QRZ-decomposition which is actually a QR factorization with column pivoting (or column reordering) [10]. In CPS MILS, the computation cost on QRZ-decomposition is lower due to the smaller dimension of $\bar{\mathbf{Q}}_A^\intercal B$, see Row (2c) in Table I. On the other hand, the computation of the integer search represented with $(*)$ in Row (2d) of Table I will not vary significantly, since the dimension of $\delta \mathbf{N}$ is the same in both approaches. The computation on the float solution and the integer validation is the same in both the Direct MILS and CPS MILS methods.

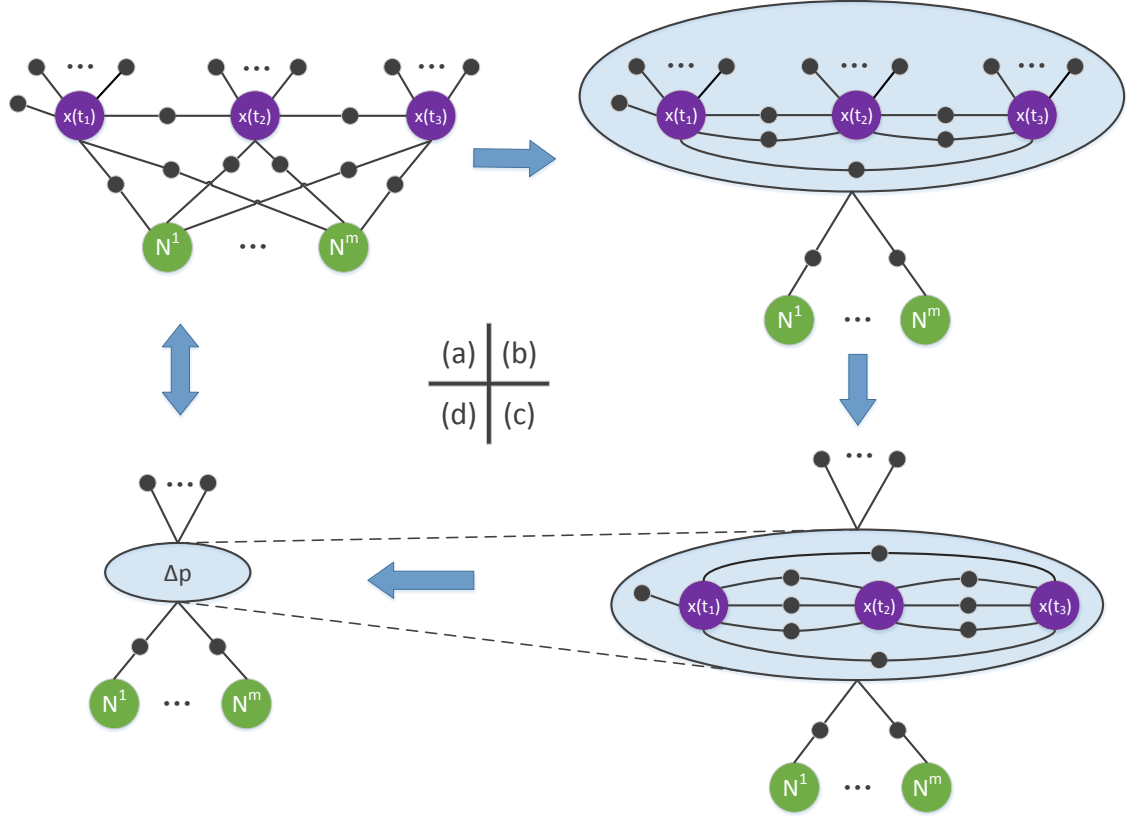### 6.3.5 Factor Graph Representation of the CPS approach

Figure 6.3: The Factor Graph representation of the CPS approach.

Fig. 6.3 shows the Factor Graph representation of the CPS approach discussed in this chapter. Fig. 6.3(a) is the factor graph shown in Section 5.3 for the original NMILS problem solved in Section 5.4. Fig. 6.3(b) indicates the process of formalizing the Integer-Free optimization. In Fig. 6.3(b), the integer ambiguity $\mathbf{N} = [N^1, \ldots, N^m]^\intercal$ is 'marginalized out' from the graph in the shaded circle, through the left-nullspace projection, see eqn. (6.11). The factors below the shaded circle

are for the processed phase measurement costs $\|\boldsymbol{r}_3(\mathbf{X}, \mathbf{N})\|^2$. By doing graph optimization of the graph inside the shaded circle which represents the cost $\|\boldsymbol{r}_a(\mathbf{X})\|^2$, the integer-free solution (equivalent to the float solution, see Proposition 6.5) can be found. Fig. 6.3(c) indicates the process of formalizing the CPS estimation. The factors above the shaded circle are for the processed code measurement costs $\|\boldsymbol{r}_2(\mathbf{X})\|^2$. Fig. 6.3(d) represents the factor graph for the CPS estimation, where the only unknown variables are the CPS $\Delta\boldsymbol{p}$ and the integer ambiguity $\mathbf{N}$. The smaller circle in 6.3(d) replaces the factor graph in the shaded circle in Fig. 6.3(d), since the latter one is independent of $(\Delta\boldsymbol{p}, \mathbf{N})$.

6.4 Implementations

This section discusses the practical implementations of the proposed CPS method. Implementation results show that the computation cost is reduced significantly by the CPS method while the state estimates are close to the original approach.

6.4.1 Discussion of CPS Implementations

In practice, besides the integer validation techniques, a threshold Δ_f for the common position shift is also used as a way for sanity check. This threshold can be picked by the designer based on the expected positioning accuracy of the float solution. After the CPS estimation in eqn. (6.9), if $\|\Delta\boldsymbol{p}^\star\| < \Delta_f$, and $\mathbf{N}^\star$ can be validated with

standard integer validation techniques [77], then the estimate of $(\mathbf{X}, \mathbf{N})$ that resulted from this common position shift approach is finalized as $(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star, \mathbf{N}^\star)$. Otherwise, the estimate $(\Delta \boldsymbol{p}^\star, \mathbf{N}^\star)$ is invalidated, and then the original full NMILS in eqn. (6.6) may be executed to get $(\mathbf{X}^*, \mathbf{N}^*)$. Alternatively, to satisfy the real-time requirement, the float solution $\check{\mathbf{X}}$ is used to update the real-time state estimates and the current CRT window could be adapted by different strategies for the next trial of integer resolution. For example, the designer can choose to slide the current window to the next epoch(s), or extend the length of the current window to accumulate more data. Or, the scheme could just skip the current window and wait for a new window with larger m, which should yield a higher success rate.

6.4.2 Experiments Description

For performance evaluation, the proposed approach is implemented in C++ and applied to post-process several RTK GPS/INS data sets collected from an automotive vehicle. The on-vehicle GPS/INS suite consists of a NovAtel OEMV3 receiver which outputs GPS pseudorange and carrier phase measurements at $1Hz$, and a $200Hz$ NV-IMU1000 Microelectromechanical system (MEMS) IMU from NAV Technology Co., Ltd. which outputs the specific forces and angular rate measurements along the three orthogonal axes. The IMU's specification sheet can be found in http://www.nav.cn/en/. The differential GPS information are from the UC Riverside Ntrip caster (ntrip.engr.ucr.edu:2101) which broadcasts raw dual frequency GPS

measurements (Message 1004 in RTCM3.1 standard) and the base position (Message 1006 in RTCM3.1 standard) publicly over the internet at $1Hz$ and $0.1Hz$, respectively. For more detials about RTCM and Ntrip protocols, see Section 2.2.2.

Two data sets logged on vehicle are post-processed by a C++ program implementing the proposed Common Position Shift approach:

1. a 12 hours (43200 seconds) stationary data collected on UC Riverside (UCR) campus, on March 29, 2014;

2. a 640 seconds moving data collected on the roads around UCR Center for Environmental Research and Technology (CE-CERT), on January 23, 2014;

All the post-processing was executed on a desktop computer with Intel Core2 Q9400 four-core CPU at 2.66GHz, 8GB DDR3 1333MHz memory, 240GB SSD disk drive. The post-processing C++ program runs on ubuntu 12.04 64-bit OS which installed within a VMware virtual machine for Windows 7. The total memory used by the virtual machine is up to 4 GB. Under this implementation environment and picking the CRT window length 10 seconds, the average computing time cost on one iteration of CPS is around 0.250 ms, versus that of the original full MILS method around 150 ms. This indicates the significant performance improvement of the CPS method on computation.

For the stationary data, the 3D position ground truth is pre-surveyed so positioning errors are presented herein to show the integer ambiguity resolution ability of

the CPS method. For the moving data sets, the ground truth is unavailable but the trajectory estimates and integer estimates from the CPS and the original full MILS method [13] are compared, to show that the CPS is an efficient alternative of the full MILS method. For both implementations, dual frequency carrier phase measurements are utilized to form wide-lane phase measurements [22]. For both stationary and moving data, the estimated residuals of L1 carrier phase measurements from each estimation are shown to illustrate the performance of the CRT estimator. The CRT Window lengths in all experiments are 10 seconds. For each available GPS satellite, a 10 deg elevation mask is applied to avoid outliers. Only the CRT windows with more than 5 satellites were tried to solve integers. To evaluate the proposed CPS method, every 10-second window is tried to do the trajectory and integer estimation, independently. In other words, for each trail of estimation with 10-second data, no estimate knowledge from other trials is used. After each estimation, a residual check is used to validate the integer estimates and avoid measurement outliers that if the magnitude of any residual of the carrier phase measurements with fixed integer estimates is larger than 0.060 meter, the integer estimates are rejected. Other integer validation techniques, e.g. ratio test, can be applied [77, 78]. Only validated results are recorded.

6.4.3 Stationary data

In the implementation with stationary data, the positioning results at each G-PS measurement time are compared with the ground truth. The survey accuracy of the ground truth is millimeter. The algorithm itself has no knowledge that the system is stationary. In particular, for each 10-second CRT windows, $\mathbf{X} = [\boldsymbol{x}(t_1)^\intercal, \boldsymbol{x}(t_2)^\intercal, \ldots, \boldsymbol{x}(t_{10})^\intercal]^\intercal$ is estimated. For the 43200-sec stationary data, 38167 trials are validated, i.e. with enough satellite numbers and passed the residual check. The positioning results are presented through comparison with the ground truth, i.e., $\boldsymbol{p}(t_k) - \boldsymbol{p}_0$ where $\boldsymbol{p}_0$ is the ground truth. To evaluate the performance of the positioning, the maximal horizontal positioning errors over each CRT window were logged. The histogram of the maximal horizontal positioning errors over the window are shown in the up-left subplot of Fig. 6.4. The up-left subplot of Fig. 6.4 shows that most of the maximal positioning errors are below 2 centimeters, which matches the expected performance of RTK GPS positioning. The down-left subplot of Fig. 6.4 shows that most of the L1 phase measurement estimated residuals lie in $[-0.020, 0.020]$.

6.4.4 Moving Data

The route and satellite availability while logging this dataset is shown in Fig. 6.5. The average vehicle speed is about 35km/h.
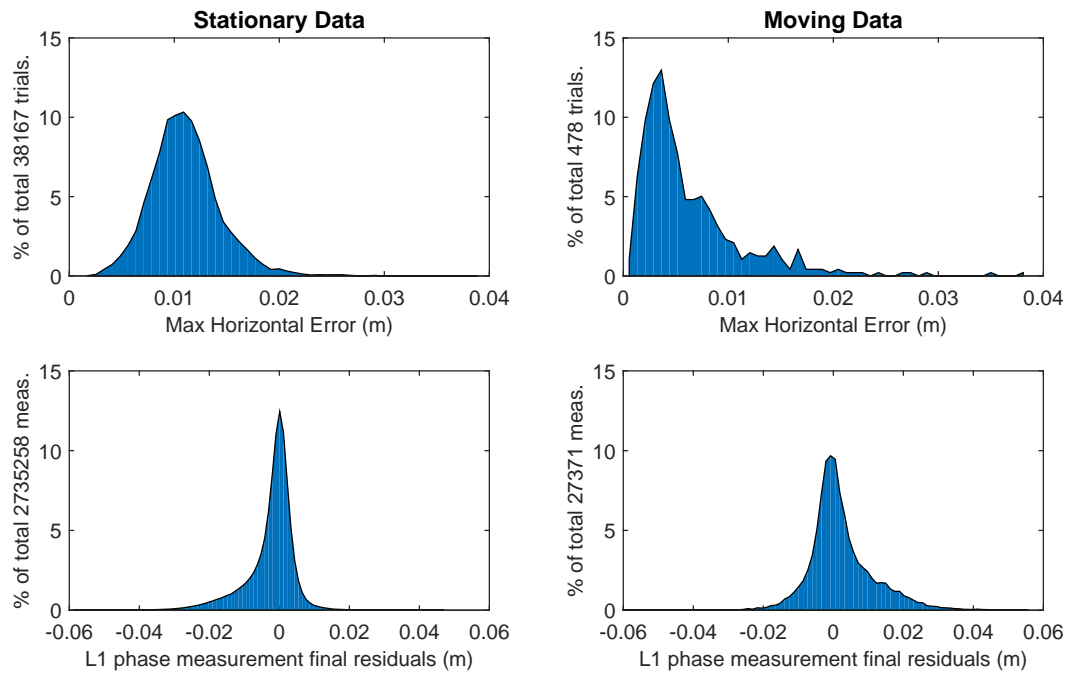
Figure 6.4: Common Position Shift method horizontal positioning errors and estimated L1 phase measurements residuals.

Figure 6.5: The route and the satellite availability of the 640s moving data experiment. The x and y axis are the longitude and latitude in degrees. The colors along the route indicate the number of satellites visible to the receiver at that location and time.

Since the ground truth is not available for this dataset, the implementation results of the CPS method are compared with those from the original full MILS method [13]. Implementation results show that all the validated integers from CPS are identical to those from the full MILS method. Furthermore, the upper-right subplot of Fig. 6.4 shows that the maximum norm, during each CRT window, of the horizontal position error between CPS and full MILS are bounded by 0.04 meters and typically less than 0.02 meters. This demonstrates that the CPS method is a good approximation of the full MILS method in terms of integer ambiguity resolution and positioning. The lower-right subplot of Fig. 6.4 shows that most of the L1 phase measurement residuals lie in $[-0.02, 0.02]$ meters.

Fig. 6.6-6.9 shows the maximal differences between the CPS estimates and the full MILS estimates of attitude, velocity, bias. Fig. 6.6 shows that the roll and pitch estimate errors are smaller than 0.1 degree and for yaw angle most of the errors are smaller than 0.5 degree. Fig. 6.7 shows that the velocity estimate errors between two methods are smaller than 0.02 m/s.
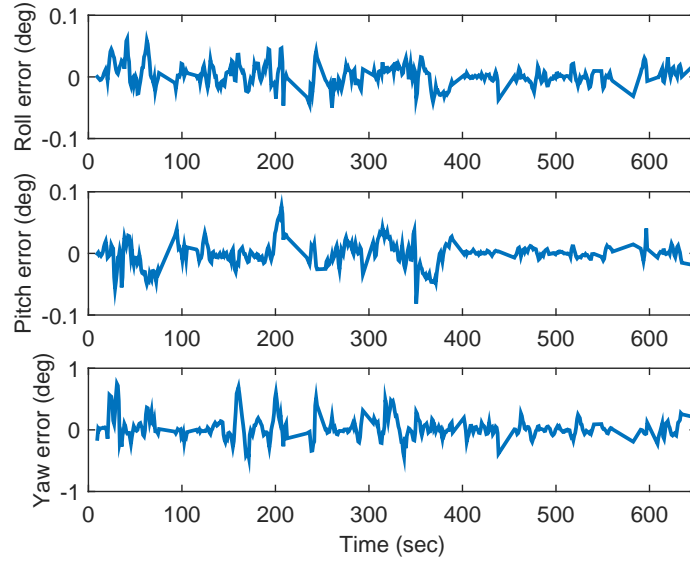

Figure 6.6: Max attitude estimation errors between CPS and full MILS results of the 640s moving data

Herein, Proposition 6.6 is validated with the experimental result. For the i-th satellite in each time window, defined

$$\Delta \boldsymbol{r}^i \triangleq \bar{\mathcal{Q}}_1^\mathsf{T} \left(\boldsymbol{h}^i \big(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star \big) - \boldsymbol{\varphi}^i \right) - \bar{\mathcal{Q}}_1^\mathsf{T} \left(\boldsymbol{h}^i \big(\check{\mathbf{X}} \big) - \boldsymbol{\varphi}^i \right)$$

as the variation of the integer-free measurement (see Proposition 6.15 or [81]) residuals caused by the estimated common-position-shift $\Delta \boldsymbol{p}^\star$. Furthermore, $\Delta \boldsymbol{r}^i$ can be
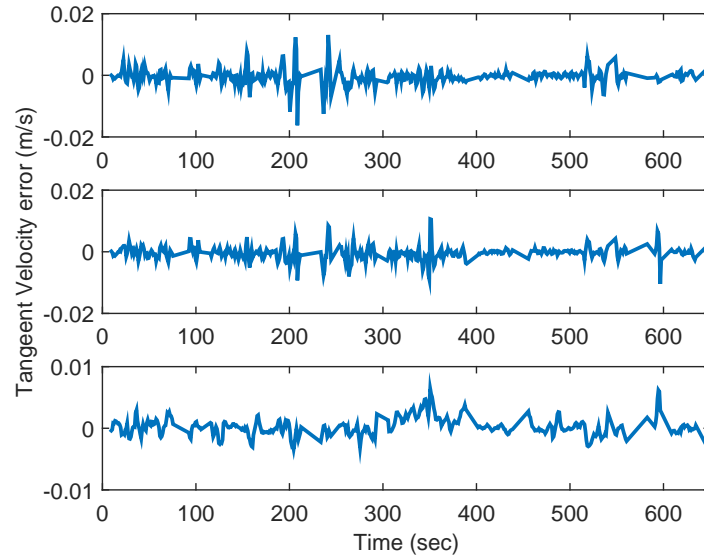
Figure 6.7: Max tangent velocity estimation errors between CPS and full MILS results of the 640s moving data
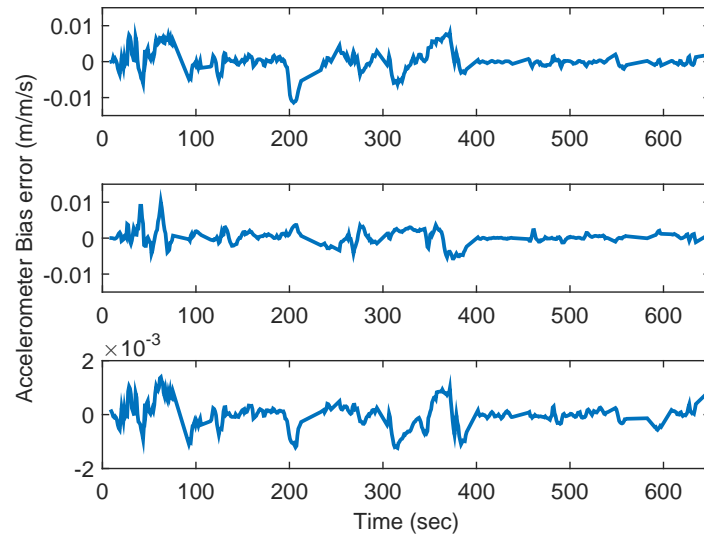

Figure 6.8: Max accelerometer bias estimation errors between CPS and full MILS results of the 640s moving data
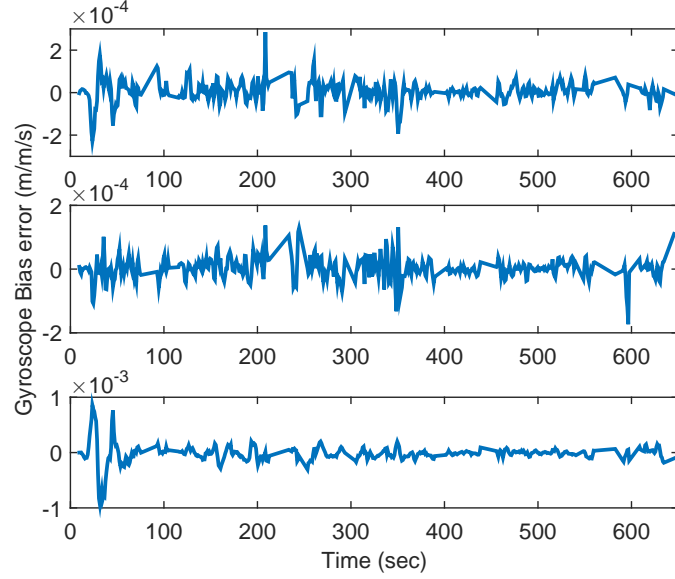
178

Figure 6.9: Max gyroscope bias estimation errors between CPS and full MILS results of the 640s moving data

rewritten as

$$\Delta \boldsymbol{r}^i = \bar{\mathcal{Q}}_1^\mathsf{T} \left(\boldsymbol{h}^i \big(\check{\mathbf{X}} \oplus \Delta \boldsymbol{p}^\star \big) - \boldsymbol{h}^i \big(\check{\mathbf{X}} \big) \right).$$

In this experiment with the moving data, the maximal magnitude of $\Delta \boldsymbol{r}^i$ over each window, i.e.

$$\boldsymbol{\delta}_{max} = \max_i \left\{ \| \Delta \boldsymbol{r}^i \|_\infty \right\}$$

is recorded along with the magnitude of common-position-shift $\| \Delta \boldsymbol{p}^\star \|$.

Proposition 6.6 implies that a common-position-shift will only cause small variations $\Delta \boldsymbol{r}^i$ and is upper-bounded by $B_1(\| \Delta \boldsymbol{p}^\star \|, \bar{v})$. Fig. 6.10 validates this claim by comparing $\boldsymbol{\delta}_{max}$, $\| \Delta \boldsymbol{p}^\star \|$ and the bound B_1 calculated in eqn. (6.22). Fig. 6 also shows that in this data set, all the common-position-shift estimates are with 1.2 meter.
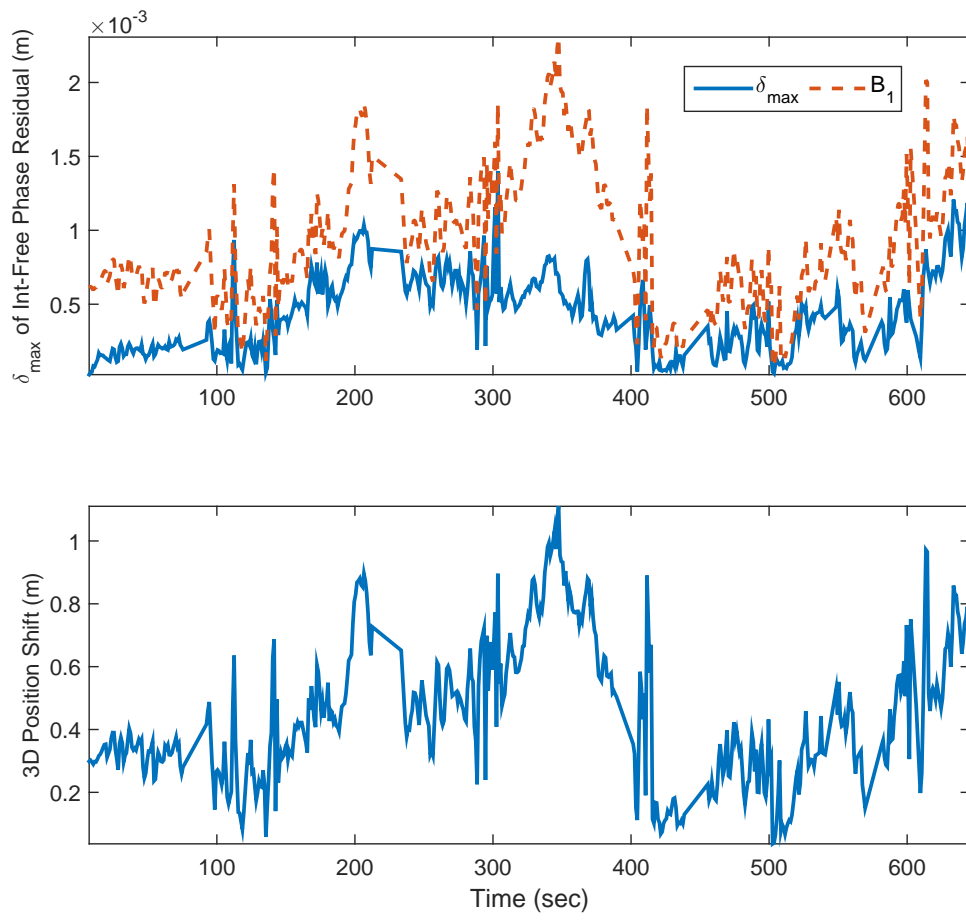
Figure 6.10: The magnitude of the 3D position shift estimated by CPS and The variations of $\|r_4\|^2$ caused by the common position shift.

Chapter 7

Conclusion and Future Works

This chapter concludes this dissertation and points out potential future works.

7.1   Conclusions

the major contribution of this dissertation is the proposition of a novel Common-Position-Shift (CPS) approach for solve integer ambiguity over a time window reliably and efficiently. This dissertation also has the following other contributions to improve the accuracy, reliability and robustness of INS, especially for land vehicle applications with GPS aiding:

- Instead of the traditional filtering frame which only considers one epoch aiding measurements, Bayesian smoothing with all navigation information over a time window is applied for INS. Then, for RTK GPS/INS case, the integer ambiguity

resolution is executed with a window of IMU and GPS measurements which provide more constraints and improve the accuracy and reliability.

- Probabilistic graphical modeling is introduced to represent the more complicated estimation problem in the smoothing for aided INS. Furthermore, Factor Graph modeling which is widely applied in robotics community, is extended for the RTK GPS/INS application by incorporating the integral unknowns to resolve integer ambiguity.

- To overcome the linearization issue inherited from EKF, iterative nonlinear optimization is applied to solve the *Maximum-a-Posteriori* estimation derived from the graphical model. Thus, previous wrong linearization points can corrected when more measurements included in the optimization of the later windows. For RTK GPS/INS case, the method of solving the Nonlinear Mixed Integer Least Square (NMILS) for smoothing is proposed.

- A novel robust optimization approach is proposed by introducing hypothesis tests to improve the existing methods. This method can be applied in graph optimization base INS, and also fit general least square problems.

- A Contemplative Real-Time (CRT) framework, which combines traditional filtering and the robust graph optimization based smoothing, is proposed for general INS and implemented especially for RTK GPS/INS.

- A novel base correction computation approach is proposed for real-time application with enhanced robustness.

7.1.1 Publication list

[1] Y. CHEN, S. ZHAO, AND J. A. FARRELL, "Computationally Efficient Carrier Integer Ambiguity Resolution in GPS/INS: a Common-Position-Shift approach," IEEE Transactions on Control Systems Technology (submitted).

[2] S. ZHAO, Y. CHEN, AND J. FARRELL, "High Precision Vehicle Navigation in Urban Environments using a MEM's IMU and Single-frequency GPS Receiver", IEEE Transactions on Intelligent Transportation Systems (submitted)

[3] Y. CHEN, S. ZHAO, D. ZHENG, AND J. A. FARRELL, "High Reliability Integer Ambiguity Resolution of 6DOF RTK GPS/INS," 53th IEEE Conference on Decision and Control, 2014, Los Angeles.

[4] S. ZHAO, Y. CHEN, AND J. A. FARRELL, "High Precision Vehicle Navigation in Urban Environment using Low-cost Single-frequency GPS Receiver," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)/Workshop on Planning, Perception and Navigation for Intelligent Vehicles (PPNIV), 2014, Chica-

go.

[5] S. ZHAO, Y. CHEN, H. ZHANG, AND J. A. FARRELL, "Differential GPS aided Inertial Navigation: a Contemplative Real-Time Approach," 19th World Congress of the International Federation of Automatic Control (IFAC), 2014, Cape Town.

[6] Y. CHEN, D. ZHENG, P. A. MILLER AND J. A. FARRELL, "Underwater Inertial Navigation with Long Base Line Transceivers: A Near-Real-Time Approach," IEEE Transactions on Control Systems Technology (minor revision).

[7] Y. CHEN, D. ZHENG, P. A. MILLER AND J. A. FARRELL, "Underwater Inertial Navigation with Long Base Line Transceivers: A Near-Real-Time Approach," 52th IEEE Conference on Decision and Control, Florence, 2013.

[8] Y. CHEN, D. ZHENG, P. A. MILLER AND J. A. FARRELL, "Underwater Vehicle Near Real Time State Estimation," IEEE Multi-Conference on Systems and Control, Hyderabad, 2013.

[9] W. DONG, Y. ZHAO, Y. CHEN AND J. A. FARRELL, "Tracking Control for Nonaffine Systems: A Self-organizing Approximation Approach," IEEE Trans. on

Neural Networks and Learning Systems, 2012.

[10] Y. CHEN, W. DONG AND J. A. FARRELL, "Locally Linearized Optimal Tracking Control for Unknown Nonlinear Systems," 50th IEEE Conference on Decision and Control and European Control Conference, 2011.

[11] Y. CHEN, W. DONG AND J. A. FARRELL, "Locally Linearized Optimal Regulation for Unknown Nonlinear Systems," IEEE Multi-Conference on Systems and Control 2011, Denver.

[12] Y. CHEN AND J. A. FARRELL, "Self-organizing Approximation Based Control with $\mathcal{L}_1$ Transient Performance Guarantees," Proceedings of American Control Conference 2011, San Francisco, Jun 29-July 1, 2011.

7.2   Future Works

The following three directions may be interesting for further investigations:

- Vision aided GPS/INS in urban canyon environment. In most of highway environment, the proposed RTK GPS/INS can realize reliable lane-level positioning. However, in urban canyon environment where the availability of satellites

are limited, other aiding measurements are highly needed. Vision sensors like camera, LIDAR are expected to help. In particular, Vision-Inertial-Odometry techniques [43–49] seems very promising to improve the GPS/INS performance in urban canyon environment.

- Expand the current model. Additional variables can be added into the current RTK GPS/INS model to improve the performance, due to the improved degree-of-freedom of the estimation. For example,

  - Receiver clock bias can also be added into the state. Currently, double difference is used to remove receiver clock bias in the estimation. Double difference processing brings correlation between the GPS measurements at the same epoch. To ensure optimality, this correlation has to be considered and would impair the original sparsity of the problem.

  - The base correction presented in Chapter 2 can also be incorporated into the Bayesian smoothing framework. Currently, the base correction computation is relatively separated routine in the framework. In future works, base measurements can also be tightly-coupled into the system, by adding the variables, which we need to estimated for base correction computation, into the state vector.

  - Multipath error modeling can also be incorporated. In this work, outliers in the GPS measurements are treated as independent blunders, without con-

sidering the correlation between them. Proper models for multipath errors could help improve the performance of the outlier handling in mutltipath prone environment.

- Extend the RTK GPS/INS to PPP GPS/INS, where PPP stands for Precise-Point-Positioning [71] which is a popular research topic in GNSS and Geodesy community. The PPP technique aims to get rid of the nearby base station but still reach centimeter level accuracy. The approaches and frameworks proposed in this work are promising to be extended to PPP GPS/INS applications and improve the state-of-the-art PPP performance.

Bibliography

[1] Online Positioning User Service: http://www.ngs.noaa.gov/opus/.

[2] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger. Closest Point Search in Lattices. *IEEE Trans Inform Theory*, 48:2201C2214, 2002.

[3] Anonymous. Global positioning system standard positioning service performance standard. Technical report, U.S. Department of Defense, 2008.

[4] W. Baarda. A Testing Procedure for Use in Geodetic Networks. *Netherlands Geod. Comm.*, 2(4), 1968.

[5] U. I. Bhatti and W. Y. Ochieng. Detecting Multiple failures in GPS/INS integrated system: A Novel architecture for Integrity Monitoring. *Journal of Global Positioning Systems*, 8(1):26–42, 2009.

[6] Stephen Boyd. *Convex Optimization*. Cambridge University Press, 2004.

[7] Robert Grover Brown and Patrick Y.C. Hwang. *Introduction to Random Signals and Applied Kalman Filtering with Matlab Exercises, 4th Edition*. Wiley, 2012.

[8] California State. Documents associated with SB 1298 in the 2011-2012 Session, Sep 2012.

[9] X. W. Chang, X. Yang, and T. Zhou. MLAMBDA: a Modified LAMBDA Method for Integer Least-Squares Estimation. *Journal of Geodesy*, 79(9):552–565, 2005.

[10] X.W. Chang and G. H. Golub. Solving Ellipsoid-Constrained Integer Least Squares Problems. *SIAM J. MATRIX ANAL. APPL.*, 31(3):1071–1089, 2009.

[11] X.W. Chang and T. Zhou. MILES: MATLAB Package for Solving Mixed Integer LEast Squares Problems. *GPS Solution*, 11(4):289–294, 2007.

[12] X. Chen, F. Dovis, S. Peng, and Y. Morton. Comparative studies of gps multipath mitigation methods performance. *IEEE Transactions on Aerospace and Electronic Systems*, 49(3):1555–1568, July 2013.

[13] Y. Chen, S. Zhao, D. Zheng, and J.A. Farrell. High Reliability Integer Ambiguity Resolution of 6DOF RTK GPS/INS. In *the Proceedings of the IEEE Conference on Decision and Control*, Los Angeles, December 2014.

[14] Y. Chen, D. Zheng, P. M. Miller, and J.A. Farrell. Underwater Inertial Navigation with Long Base Line Transceivers: A Near-Real-Time Approach. In *the Proceedings of the IEEE Conference on Decision and Control*, pages 5042–5047, Florence, Italy, December 2013.

[15] Y. Chen, D. Zheng, P. M. Miller, and J.A. Farrell. Underwater Vehicle Near Real Time State Estimation. In *the Proceedings of the IEEE Multi-Conference on Systems and Control*, pages 545–550, Hyderabad, India, August 2013.

[16] Y. Chen, D. Zheng, P. M. Miller, and J.A. Farrell. Underwater Inertial Navigation with Long Base Line Transceivers: A Near-Real-Time Approach. *IEEE Trans. on Control Systems Technology (accepted)*, 2014.

[17] F. Dellaert. Factor Graphs and GTSAM: A Hands-on Introduction. Technical report, Georgia Institute of Technology, 2012.

[18] Frank Dellaert and Michael Kaess. Square Root SAM: Simultaneous Localization and Mapping via Square Root Information Smoothing. *Int. J. Rob. Res.*, 25(12):1181–1203, 2006.

[19] D. Dettmering and G. Weber. The euref-ip ntripbroadcaster: Real-time gnss data for europe. In *Proceedings of the IGS Workshop*, pages 1–8, Astronimical Institute University of Bern, Switzerland, 2004.

[20] Tue-Cuong Dong-Si and Anastasios I Mourikis. Motion Tracking with Fixed-lag Smoothing: Algorithm and Consistency Analysis. In *IEEE ICRA*, pages 5655–5662, 2011.

[21] J.A. Farrell, T.D. Givargis, and M.J. Barth. Real-time Differential Carrier Phase GPS-aided INS. *IEEE Trans. CST*, 8(4):709–721, 2000.

[22] Jay A. Farrell. *Aided Navigation: GPS with High Rate Sensors*. McGraw Hill, 2008.

[23] Udo Frese. A Proof for the Approximate Sparsity of SLAM Information Matrices. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, 2005.

189

[24] Y. Gao, E. J. Krakiwsky, and J. Czompo. Robust Testing Procedure for Detection of Multiple Blunders. *Journal of Surveying Engineering*, 118, 1992.

[25] Gene H Golub and Charles F Van Loan. *Matrix Computations (3rd Edition)*. The Johns Hopkins University Press, 1996.

[26] T. Hastie, R. Tibshirami, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer, 2nd edition, 2009.

[27] S. Hewitson and J Wang. Extended Receiver Autonomous Integrity Monitoring (eRAIM) for GNSS/INS Integration. *Journal of Surveying Eng.*, 136(1):13–22, 2010.

[28] Steve Hewitson, Hung Kyu Lee, and Jinling Wang. Localizability Analysis for G-PS/Galileo Receiver Autonomous Integrity Monitoring. *J. of Navigation*, 57:245–259, 2004.

[29] Steve Hewitson and Jinling Wang. GNSS Receiver Autonomous Integrity Monitoring (RAIM) Performance Analysis. *GPS Solution*, 10:155–170, 2006.

[30] D. Hoaglin, F. Mosteller, and J. Tukey. *Understanding Robust and Exploratory Data Analysis.* John Wiley and Sons, Inc., New York, 1983.

[31] Peter J. Huber. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 1964.

[32] Peter J. Huber. *Robust Statistics.* John Wiley and Sons, New York, 1981.

[33] P. Jonge and C Tiberius. The LAMBDA Method for Integer Ambiguity Estimation: Implementation Aspects. Technical report, TUDelft, 1996.

[34] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J.J. Leonard, and F. Dellaert. iSAM2: Incremental smoothing and mapping using the Bayes tree. *Intl. J. of Robotics Research*, 31(2):216–235, February 2012.

[35] M. Kaess, A. Ranganathan, and F. Dellaert. iSAM: Incremental Smoothing and Mapping. *IEEE Trans. on Robotics*, 24(6):1365–1378, 2008.

[36] E.D. Kaplan and C.J. Hegarty. *Understanding GPS Principles and Applications, 2nd Ed.* Artch House, 2006.

[37] Steven M. Kay. *Fundamentals of Statistical Signal Processing, Estimation theory.* Prentice Hall PTR, 1993.

[38] D. Koller and N. Friedman. *Parameter Estimation and Hypothesis Testing in Linear Models.* Spring-Verlag, New York, N.Y., 1988.

[39] D. Koller and N. Friedman. *Probabilistic Graphical Models*. The MIT Press, 2009.

[40] F.R. Kschischang, B.J. Frey, and H.-A. Loeliger. Factor Graphs and the Sum-product Algorithm. *Information Theory, IEEE Transactions on*, 47(2), 2001.

[41] Rainer Kuemmerle, Giorgio Grisetti, Hauke Strasdat, Kurt Konolige, and Wolfram Burgard. g2o: a General Framework for Graph Optimization. In *Proc. ICRA*, 2011.

[42] A.K. Lenstra, H.W. Lenstra, and Lovasz L. Factoring polynomials with rational coefficients. *Math Ann*, 261:515–534, 1982.

[43] M. Li, B.H. Kim, and A. I. Mourikis. Real-time motion estimation on a cellphone using inertial sensing and a rolling-shutter camera. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 4697–4704, Karlsruhe, Germany, May 2013.

[44] M. Li and A. I. Mourikis. High-precision, consistent EKF-based visual-inertial odometry. *International Journal of Robotics Research*, 32(6):690–711, May 2013.

[45] M. Li, H. Yu, X. Zheng, and A. I. Mourikis. High-fidelity sensor modeling and self-calibration in vision-aided inertial navigation. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 409–416, Hong Kong, June 2014.

[46] Mingyang Li and A. I. Mourikis. Optimization-based estimator design for vision-aided inertial navigation. In *Proceedings of Robotics: Science and Systems*, Sydney, Australia, Jul. 2012.

[47] Mingyang Li and Anastasios I. Mourikis. Improving the accuracy of EKF-based visual-inertial odometry. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 828–835, St Paul, MN, May 2012.

[48] Mingyang Li and Anastasios I. Mourikis. Vision-aided inertial navigation for resource-constrained systems. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1057–1063, Vilamoura, Portugal, Oct. 2012.

[49] Mingyang Li and Anastasios I Mourikis. Online temporal calibration for camera-IMU systems: Theory and algorithms. *The International Journal of Robotics Research*, 33(6), June 2014.

[50] H.-A. Loeliger. An Introduction to Factor Graphs. *IEEE Signal Processing Magazine*, 21(1):28–41, 2004.

[51] Yi Ma, Jana Košecká, Stefano Soatto, and Shankar Sastry. *An Invitation to 3-D Vision From Images to Models.* Springer, New York, 2004.

[52] E. M. Mikhail and F. E. Acherman. *Observations and Least Squares.* Univ Pr of Amer, 1976.

[53] P.A. Miller, J.A. Farrell, Y. Zhao, and V. Djapic. Autonomous Underwater Vehicle Navigation. *IEEE J. of Oceanic Eng.*, 35(3):663 – 678, 2010.

[54] P. Misra and P. Enge. *GPS: Signals, Measurements, and Performance.* Ganga-Jamma Press, 2001.

[55] A.I. Mourikis and S.I. Roumeliotis. A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation. In *IEEE ICRA*, pages 3565–3572, 2007.

[56] Gary P. and C. E. Fly. NDGPS assessement final report. Technical report, U.S. Department of Transportation, 2008.

[57] Gary P. and C. E. Fly. RTCM Standard 10410.1, Networked Transport of RTCM via Internet Protocol (Ntrip), Version 2.0. Technical report, Radio Technical Commission for Maritime Services, 2011.

[58] D. M. Rosen, M. Kaess, and J. Leonard. An Incremental Trust-Region Method for Robust Online Sparse Least-Squares Estimation. In *IEEE ICRA*, 2012.

[59] RTCM Special Committee No. 104. RTCM Standard 10403.1, Differential GNSS Services, Version 3. Technical report, Radio Technical Commission for Maritime Services, 2011.

[60] R. A. Snay and M.ASCE Tomás Soler. Continuously operating reference station (CORS): History, applications, and future enhancements. *Journal of Surveying Engineering*, 134(4), 2008.

[61] Stanford University. Automated driving: Legislative and regulatory action, http://cyberlaw.stanford.edu/wiki. Online.

[62] N. Sünderhauf, M. Obst, G. Wanielik, and P. Protzel. Multipath Mitigation in GNSS-based Localization using Robust Optimization. In *Proc. of IEEE Intelligent Vehicles Symposium*, 2012.

[63] Swift Navigation. Piksi is a low-cost, high-accuracy gps receiver with real time kinematics (rtk): http://www.swiftnav.com/.

[64] T. Takasu and A. Yasuda. Evaluation of RTK-GPS performance with Low-cost Single-frequency GPS Receivers. *Proceedings of International Symposium on GPS/GNSS*, pages 852–861, 2008.

[65] Tomoji Takasu and Akio Yasuda. Development of the Low-cost RTK-GPS Receiver with an Open Source Program Package RTKLIB. In *International Symposium on GPS/GNSS, International Convention Center Jeju, Korea*, 2009.

[66] P. J. G. Teunissen. Least-Squares Estimation of the Integer GPS Ambiguities. In *Invited lecture, Section IV: Theory and Methodology*. IAG General Meeting, Beijing, China., 1993.

[67] P. J. G. Teunissen and A. Kleusberg. *GPS for Geodesy*. Springer, 2nd edition, 2005.

[68] P J G Teunissen and S Verhagen. THE GNSS AMBIGUITY RATIO-TEST REVISITED : A BETTER WAY OF USING IT. *Survey Review*, 151(April):138–151, 2009.

[69] P.J.G Teunissen. The Least-Squares Ambiguity Decorrelation Adjustment: a Method for Fast GPS Ambiguity Estitmation. *Journal of Geodesy*, 70:65–82, 1995.

[70] P.J.G Teunissen. An Optimality Property of the Integer Least-Squares Estimator. *Journal of Geodesy*, 73:587–593, 1999.

[71] P.J.G Teunissen, D. Odijk, and B. Zhang. PPP-RTK: Results of CORS Network-Based PPP with Integer Ambiguity Resolution. *Journal of Aeronautics, Astronautics and Aviation*, 42(4), 2010.

[72] S. Thrun, W. Burgard, and Dieter Fox. *Probabilistic Robotics*. The MIT Press, 2006.

[73] Robert Tibshirami. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.

[74] Nikolas Trawny and Stergios I Roumeliotis. Indirect Kalman Filter for 6D Pose Estimation. Technical report, Multiple Autonomous Robotic Systems Laboratory, University of Minnesota, Minneapolis, MN, 2005.

[75] S. Verhagen, D. Odijk, P.J.G. Teunissen, and L. Huisman. Performance improvement with low-cost multi-gnss receivers. In *Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC), 2010 5th ESA Workshop on*, pages 1–8, 2010.

[76] S. Verhagen, P.J.G. Teunissen, and D. Odijk. The future of single-frequency integer ambiguity resolution. In *VII Hotine-Marussi Symposium on Mathematical Geodesy*, volume 137 of *International Association of Geodesy Symposia*, pages 33–38. Springer Berlin Heidelberg, 2012.

[77] Sandra Verhagen. *The GNSS Integer Ambiguities: Estimation and Validation.* NCG, Nederlandse Commissie voor Geodesie, Netherlands Geodetic Commission, 2005.

[78] Anh Vu, J Farrell, and Matthew Barth. Centimeter-Accuracy Smoothed Vehicle Trajectory Estimation. *Intelligent Transportation Systems Magazine, IEEE,* 5(4):121–135, 2013.

[79] R. Wagner, O. Birbach, and U. Frese. Rapid Development of Manifold-based Graph Optimization Systems for Multi-Sensor Calibration and SLAM. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 3305–3312, 2011.

[80] Dong Wang, Huchuan Lu, and Ming-Hsuan Yang. Least Soft-Threshold Squares Tracking. *2013 IEEE Conference on Computer Vision and Pattern Recognition,* pages 2371–2378, June 2013.

[81] S. Zhao, Y. Chen, and J.A. Farrell. High Precision 6DOF Vehicle Navigation in Urban Environments using a Low-cost Single-frequency GPS Receiver. In *the Proceedings of IROS/PPNIV*, Chicago, September 2014.

[82] S. Zhao, Y. Chen, H. Zhang, and J.A. Farrell. Differential GPS Aided Inertial Navigation: a Contemplative Realtime Approach. In *the Proceedings of the 19th IFAC World Congress*, Cape Town, South Africa, August 2014.

[83] T. Zhou. Modified LLL Algorithms. Master's thesis, McGill University, 2006.