UCGE Reports
Number 20256

Department of Geomatics Engineering

Precise GPS Relative Navigation for Future Geodetic Satellite Formation Flying Missions

(URL: http://www.geomatics.ucalgary.ca/research/publications/GradTheses.html)

by

Raymond Tsoi

July 2007

THE UNIVERSITY OF CALGARY

Precise GPS Relative Navigation For Future Geodetic Satellite Formation Flying Missions

by

Raymond Tsoi

A DISSERTATION
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA

July, 2007

Abstract

The technique of formation flying has gained much attention in the scientific community, as it is believed to be the dominant approach for future space mission to enhance the overall performance. The added advantages of formation flying over the conventional single satellite approach include improve mission reliability with backup satellites, provide stereo observation and redundancy, and reduce the mission operation cost. Many organizations have already planned to utilize such an incredible technique for their future missions, however the use of satellite formation flying is still very new and much work is need to improve and verify the needed technologies for such an approach.

There are many important technologies associated with formation flying that would make it applicable to future space mission. One of such technology is the instrument for providing precise relative navigation. Accurate relative information is crucial for controlling the configuration to prevent collision and drift. For Low Earth Orbit mission, GPS is the ideal navigation sensor as it has proven to be reliable and precise for terrestrial applications. In this research, the navigation algorithm uses a Kalman filter that incorporates the Hill equations of motion as the propagation model. Software simulation analyzes have been conducted to verify the performance in terms of accuracies and robustness during signal outage. Although the Hill equations are only a reduced dynamic model, but promising navigation performance are delivered.

Acknowledgements

The following is a list of people who I want to sincerely thank them for their help, guidance and especially their patients towards this thesis. This work would not have been possible without their continuous support and encouragement.

Dr. Elizabeth Cannon

Dr. Nico Sneeuw

Mom, Dad & Grandma

Vincent & William

Diana & Doey

Colleagues & Friends

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Satellite technology has been a significant gift to the scientific community. Prior to the discovery of satellite technology, exploring the deep secrets of our living planet and the space beyond was only a dream; however, with the invention of satellites, not only is this dream becoming a reality, the almost endless potential of satellites is inspiring scientists to develop even more methods and approaches in experimentation to explore the untouched parts of the universe. Over the years, countless missions have been organized and launched, initially helping human kind to gain knowledge about the Earth, then the solar system and galaxy, and now helping to proceed onto discovering the secrets of the universe. Satellite technology has demonstrated its ability to allow humans to observe planets continuously. Furthermore, refinement in related technologies has helped to accelerate the gathering of scientific information and hence it has helped to open many doors, leading to new fields of science. Recent discoveries and research on gravity field recovery, satellite-based navigation and atmospheric occultation are only some examples of the frontier scientific studies and research conducted today with the help of this technology.

In the past, most conventional space missions consisted of a single satellite designed to accomplish the mission objective. The performance of these monolithic satellite missions was generally constrained by the physical dimension of the spacecraft, for example, the Hubble Space Telescope represented the limit in space image

resolution due to its single satellite approach (Leitner et al., 2002). The only way to enhance the imaging performance of the telescope was to enlarge its size but that had the consequence of increasing the production cost. With budget constraints on missions, super-sized satellites were not seen as cost effective therefore they were not considered practical solutions. In an attempt to overcome this problem, the technique of *formation flying* was proposed.

In 2000, the idea of applying formation flying into space missions was made a reality by the development of the New Millennium Program's Earth Observing-1 (EO-1) mission developed by the National Aeronautics and Space Administration (NASA). As the first autonomous formation flying mission, this program demonstrated the ability of satellites to react with one another, and simultaneously, maintain formation without constant human intervention, hence helping satellite formation flying technology secure a newfound importance in the area of satellite technology. Prior to the EO-1 mission, there were few considerations available for formation flying missions, but currently over 30 space missions have been planned utilizing this technique for exploration around the Earth and deep space (Bauer et al., 1999; Leitner et al., 2002). The advantage associated with using groups of satellites is that they permit scientists to combine data from multiple autonomous satellites. By obtaining data from multiple satellites rather than a single satellite, researchers can cut back on satellite production cost, mission risk, and operation cost but nonetheless increase the quality and quantity of information.

The many advantages associated with satellite formation flying technology have created an increase in its popularity amongst scientists; consequently, further research has been proposed and undertaken to enlighten the capabilities of future

space missions. The major goal of my research is to validate and enhance the critical technologies required for formation flying missions which can verify the possibility for future satellite formation flying missions.

1.1 Background

In 2002, NASA and the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt, or DLR) successfully launched the Gravity Recovery And Climate Experiment (GRACE) mission, the first mission designed to utilize formation flying to measure the Earth's gravity field, into the Earth's orbit from Russia's Plesetsk Cosmodrome. The orbit of GRACE was specially designed to achieve distinct characteristics for the mission. The satellites were placed in Low Earth Orbit (LEO) to increase their sensitivity to the Earth's gravity signal which in turn improved the recovery of high frequency anomalies. A high orbit inclination of 89° gave the mission near global coverage while leaving two insignificant gaps at each pole. The small eccentricity created a nearly circular orbit which makes the information collected more homogeneous and uniform. As for the relative orbit, the two identical satellites were placed in the same orbital path with one trailing the other with a mean inter-satellite distance of 220 km. Such a configuration is referred to as the leader-follower formation and it enabled the GRACE mission to measure the north-south gravitational signature of the Earth. By measuring the constantly changing distance between the two satellites with the K-band ranging system, it was possible to construct a precise map of the Earth's gravity.

Although GRACE was able to provide significant knowledge on the current geoid,

there are still limitations. The current GRACE mission measures the gradiometric along-track component of the gravity signal, which is considered to be the weaker component. The gravity field recovery capability can be strengthened by additional measurement components in the radial direction. Temporal aliasing of the high frequency signal into a monthly solution is another limitation of GRACE, but that can be avoided by including information from the across-track component. Realizing these constraints, Sneeuw (2002b) proposed the application of more complex configurations to future gravity field recovery missions. These formation types suggested are designed to overcome the limitations in the current mission. Both the EO-1 and GRACE missions have proven the applicability of the leader-follower formation for Earth observing purposes, but the idea of using more complex configurations has not yet been validated. Further research must therefore be carried out to verify the applicability of more complex configuration types to future missions.

In order to understand the relative motion of satellites, it is important to understand the equations of motion that describe the relative orbit dynamics. One such dynamic model is the linearized Hill's equations of motion (Hill, 1878; Clohessy & Wiltshire, 1960), which have been used consistently to analyze relative motion between satellites (Xu et al., 2004). In this model, the shape of the Earth is approximated by a sphere, but in reality the Earth is closer to an ellipsoid. This spherical approximation causes the dominant orbit perturbation effect known as J_2. The main effect of J_2 is the causation of secular drift in certain orbit elements (Xu et al., 2005). Any differential perturbation between satellites in the configuration can alter the formation, causing satellites to drift apart over time or even cause them to collide into one another. Without understanding the effects of the spherical approximation, and

without considering the considerations of the J_2 perturbations when designing the configurations, additional fuel will be required in order to compensate for the differential secular drift hence increasing production costs. Therefore, it is extremely important to design configurations that are J_2 invariant.

Another crucial requirement for formation flying is the knowledge of the *relative* state between satellites. This information must be highly accurate in order to maintain constellation control and to avoid collision. For LEO missions, the Global Positioning System (GPS) can fulfill this requirement as satellites that are in these low orbits are within the GPS constellations. GPS receivers are relatively low in cost compared to other navigation sensors and their performance is generally more reliable and readily available (Zimmerman & Cannon Jr., 1995; Adams et al., 1996; Olsen et al., 1998). Realizing such advantages, GPS is highly influential to the success of formation flying for future satellite missions (Bauer et al., 1999; Leitner et al., 2002). Corazzini et al. (1997) have demonstrated that the use of the differential technique, based on the carrier phase observable, is applicable for spacecraft formation flying. Ebinuma (2001) demonstrated relative positioning for two spacecraft using the double difference measurement without attempting to fix carrier phase ambiguities, which generally holds the key to consistent precise positioning. Leung & Montenbruck (2005) have also done similar research and attempted to fix the carrier phase ambiguities to integer values, the result being centimetre level positioning accuracies. Relative positioning accuracy at similar accuracy levels has also been demonstrated by the work of Busse et al. (2002b) and Park et al. (2000). The most superior performance has been demonstrated by Kroes (2006), where millimetre accuracy was achieved for the GRACE mission, however, no analysis was conducted

on more complex formation types. At the University of Calgary, several research projects have been conducted to develop high accuracy relative positioning systems. This research has been limited to land-based, shipborne and airborne applications with the reference receiver required to remain in static mode (Olynik, 2002). Crawford (2005) did demonstrate relative positioning with a moving reference station but the analysis was limited to land-based robotic vehicles. In order to extend the applications to a spaceborne environment, more work needs to be conducted.

Furthermore, GPS-based navigation in real-time usually incorporates dynamic models to refine the relative states. There are numerous dynamic models available which capture the perturbing effect of J_2. Vadali et al. (2000) presented a set of linear nonhomogeneous differential equations to describe the relative motion between satellites which includes the perturbation effect caused by J_2. Schweighart & Sedwick (2001), using a different approach, derived analytical solutions to a modified version of Hill's equations that incorporates the effect of J_2. The solutions have been further refined by Schweighart & Sedwick (2002) and simulation results indicate that the new solution set is accurate to the sub-metre level. A transition matrix has been developed by Alfriend et al. (2000b) and Gim & Alfriend (2001) that defines the change of orbital elements by considering the geometry problem and not using differential equations. The transition matrix not only accounts for the J_2 perturbation effect but it also takes into consideration the eccentricity of the reference orbit. Although these improved equations of motion can capture the higher order perturbation effect, their complexity is increased dramatically. For GPS navigation, where the sampling frequency can reach 1 Hz or higher, the use of such complex dynamic models does not provide adequate improvements in the navigation solution. Therefore the nature of

the dynamic model can be simple and well-defined for real-time navigation purposes.

1.2 Objectives

The technique of formation flying is the key tool for future space missions. This technique enables satellites to collect a greater amount of scientific information than before but at a much lower production cost. In order to apply the technique of formation flying to future satellite missions, the issues associated with it must be investigated and verified. These issues include: modelling the relative motion of satellites, investigating possible design of J_2 invariant configurations and demonstrating precise relative navigation for the formation. The objective of my research is to demonstrate the feasibility of formation flying to future space missions. The focus of this research includes the following:

- Implement the equations of motion to analyze satellite relative motion in the presence of J_2 perturbations;
- Investigate possible J_2 invariant formations that can be utilized for future geodetic missions;
- Access the accuracy of possible relative dynamic models that may be used to facilitate onboard relative navigation ;
- Quantify the typical GPS environment experienced by LEO spacecraft at 500 km altitude;
- Replicate the GPS measurements (code and carrier) and the associated errors

(orbital, ionosphere, multipath, noise) using software simulation;

- Develop real-time navigation algorithms that can deliver relative positioning at the centimetre accuracy level using GPS; and
- Assess the accuracy of the navigation algorithms in simulation mode and differentiate the performance between kinematic and dynamic models.

To complete the objectives specified, significant enhancement have been made to the GPS processing software, FLYKIN+^{TM}, developed at the University of Calgary. The modified software is now capable of processing measurements from moving base stations which allows research to extend past the Earth's surface and into spaceborne environments.

1.3 Outline

Chapter 2 describes the fundamentals related to satellite orbits; such topics include coordinate systems, coordinate transformations and the two-body problem. Several basic equations of motion are presented to describe the satellite dynamics in both homogeneous and J_2 disturbed gravitational fields.

Chapter 3 derives the linearized Hill Equations of motion (HE) which describe satellite dynamics within a local orbiting frame. Homogeneous and heterogeneous solutions for the HE are presented and the accuracy of these solutions is assessed. Based on the HE, two formation flying configuration types are examined for applicability to future space missions.

Chapter 4 describes the general measurements used for GPS navigation and discusses the characteristics of each associated error source. General observation equations for both pseudorange and carrier phase measurements are presented to illustrate how they are contaminated by different errors. Furthermore, the algorithms used in the software GPS measurement simulator, SIMGNSSII^{TM}, are thoroughly described in Chapter 5.

Chapter 6 presents the algorithms used in the GPS processing software FLYKIN+^{TM} for real-time relative navigation. Differential GPS concepts and the advantages associated are used along with the basic architecture of the Kalman filter. Two propagation models based on the kinematics and dynamics of satellites are introduced in this chapter.

In Chapter 7, the orbital trajectory, configuration and GPS environments of each test scenario are described in detail. The test methodology and parameters used to assess the relative navigation performance in the position, velocity, ambiguity domains are presented.

Chapter 8 concludes the thesis and provides further research recommendations and directions.

Chapter 2

Overview of Satellite Orbits

Understanding of satellite dynamics is essential for orbit determination. In this chapter, the basic differential equations are illustrated to study satellite motion in both central and disturbed gravitational fields. This knowledge is crucial for studying the relative motion of satellites as discussed in later chapters. Detailed description of satellite dynamics can also be found in the literature e.g. Montenbruck & Gill (2000), Schaub & Junkins (2003) and Seeber (2003).

2.1 Coordinate Systems

Prior to proceeding onto any detailed derivations or analysis, the coordinate systems used throughout this thesis are explained and clarified. The definitions of these coordinate systems are crucial for studying satellite motion and GPS navigation.

Inertial Coordinate System

The first coordinate system introduced is the inertial frame. In such a coordinate system, objects at rest tend to stay at rest and objects in motion tend to stay in motion with the same speed and direction unless acted upon by another force, which complies with Newton's 1^{st} law of motion (Newton, 1687). Such a stable coordinate system is essential for studying the dynamics of satellites as it is non-rotating and non-accelerating. As for the origin, it is located at the Earth's centre of mass, with

the $\mathbf{x}_i$ vector pointing in the direction of the vernal equinox, Υ, and the $\mathbf{z}_i$ vector pointing to the Celestial Ephemeris Pole (CEP). The $\mathbf{y}_i$ vector is perpendicular to the other axes forming a right-handed system, see Figure 2.1. All quantities referring to the inertial frame are denoted with the subscript i. Ω, I and ω are orbital parameters that describes the orientation of other coordinate systems. These parameters are thoroughly described in Section 2.2.3.

Orbital Coordinate System

The satellite trajectory can be defined by six Keplerian elements (Section 2.2.3), which are orbital parameters referenced to the inertial frame. Conversion of the orbit elements to Cartesian coordinates is not straightforward and requires the orbital coordinate system to simplify the procedure. In the orbital frame, the satellite position and velocity are expressed on the Kepler ellipse. The origin is located at the focus of the ellipse with the $\mathbf{x}_o$ vector pointing towards the perigee, the $\mathbf{z}_o$ vector being normal and pointing out from the orbital plane, and the $\mathbf{y}_o$ vector being orthogonal to the other axes forming a left-handed coordinate system, see Figure 2.1. All quantities referring to the orbital frame are denoted with the subscript o.

Local Orbiting Coordinate System

This local orbiting triad, also known as the Hill frame, is a right-handed rotating coordinate system with its origin located at the reference satellite; although in reality this origin does not have to be occupied by a physical satellite. The importance of the Hill frame is its relation with the Hill equations of motion for studying the dynamics of satellites, see Chapter 3. The orientation of this triad is defined with the $\mathbf{x}_h$ vector pointing in the radial direction, the $\mathbf{y}_h$ vector aligned to the velocity vector and the

$\mathbf{z}_h$ vector pointing in the cross-track direction forming a right-handed coordinate system, see Figure 2.1. All quantities referring to the Hill frame are denoted with the subscript h.

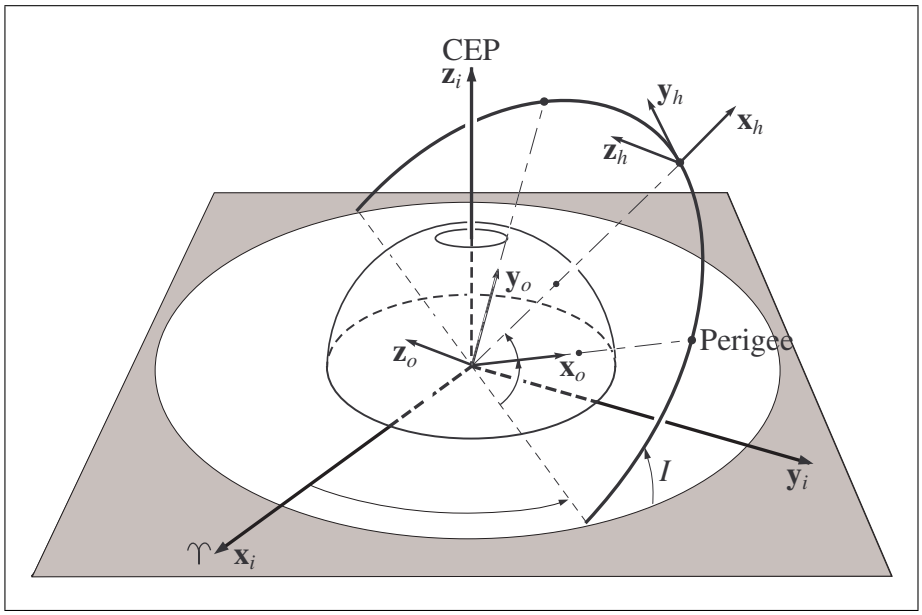


Figure 2.1: *Inertial, orbital and Hill coordinate systems (Schwarz, 1999).*

Geodetic Coordinate System

The geodetic frame is a rotating coordinate system with its origin located at the Earth's centre of mass. The $\mathbf{x}_g$ vector is the intersection between the equatorial plane and the mean meridian plane of Greenwich, the $\mathbf{z}_g$ vector points to the Conventional Terrestrial Pole (CTP) and the $\mathbf{y}_g$ vector is orthogonal to the other axes forming a right-handed coordinate system. All quantities referring to the geodetic frame are denoted with the index g.

Curvilinear measures can be used for representation of geodetic coordinates but require the definition of a reference ellipsoid. The datum ellipsoid adopted by GPS

is the World Geodetic System 1984 (WGS 84) developed by U.S. Department of Defense. The geodetic latitude φ measures the angle in the meridian plane from equator to the line perpendicular to the ellipsoid at point P. λ is the longitude which measures the angle in the equatorial plane between the Greenwich meridian and meridian plane through point P. Finally, h is the measure of the geodetic height along the ellipsoid normal, see Figure 2.2.

Local Geodetic Coordinate System

With the datum ellipsoid defined, a local geodetic coordinate system can be established. Such a coordinate system is practical for terrestrial relative positioning, as the orientation of the coordinate system is similar to the user. Relative measures can be used to define a location reference to point P which is the origin of the local geodetic frame. The $\mathbf{x}_l$ vector is tangent to the geodetic meridian pointing north, the $\mathbf{z}_l$ vector is orthogonal to theellipsoid at point P, and $\mathbf{y}_l$ points east forming a left-handed system, see Figure 2.2. All quantities referring to the local geodetic frame are denoted with the subscript l.

2.1.1 Coordinate Transformation

With all the coordinate systems defined, there is also a need to relate the various frames, which requires the formulation of coordinate transformations. The properties of reflection and rotation matrices are fundamental for coordinate transformations. Formulation of the coordinate transformation can be greatly simplified with the use of these matrices.

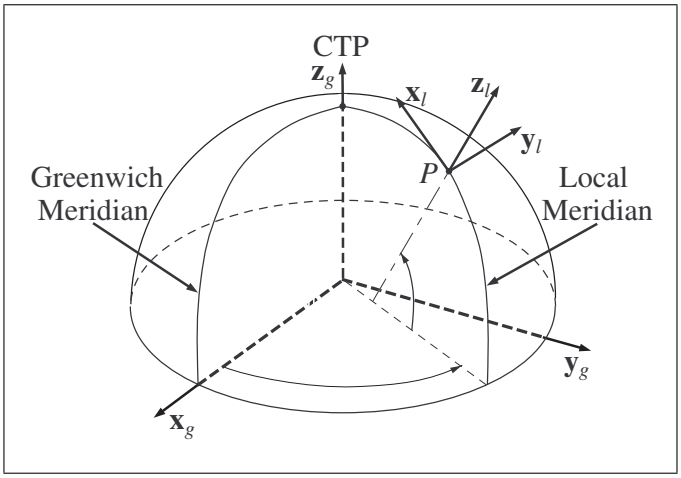


Figure 2.2: *Global and local geodetic coordinate systems (Moritz, 1980).*

Reflection Matrices

The purpose of reflection matrices is to interchange the positive and negative directions of an axis.

$$\mathbf{P}_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.1}$$

$\mathbf{P}_1$ shown is the reflection matrix that changes the sign of the x-axis and leaves the other axes unchanged. The operation reflects the coordinates about the x-axis. Similarly, the other reflection matrices are

$$\mathbf{P}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.2}$$

and

$$\mathbf{P}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \tag{2.3}$$

where $\mathbf{P}_2$ and $\mathbf{P}_3$ are for reflection in the y-axis and z-axis, correspondingly.

Rotation Matrices

The purpose of rotation matrices is the rotation of the coordinates about an axis with a certain angle α.

$$\mathbf{R}_1(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix} \tag{2.4}$$

The 3-by-3 matrix shown is the rotation matrix $\mathbf{R}_1(\alpha)$, which describes a rotation of angle α about the **x**-axis. Similarly, the rotation matrices about the two other axes are

$$\mathbf{R}_2(\alpha) = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix} \tag{2.5}$$

and

$$\mathbf{R}_3(\alpha) = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.6}$$

where $\mathbf{R}_2$ and $\mathbf{R}_3$ are the rotation matrices about the **y**-axis and **z**-axis, respectively.

Now let $\mathbf{r}_* = [x_*, y_*, z_*]^\mathsf{T}$ be the position vector in the denoted coordinate system. Using the fundamental matrices defined, the coordinate transformations between the different frames are as follows:

Orbital to Inertial (Seeber, 2003)

$$\mathbf{r}_i = \mathbf{R}_3(-\Omega)\mathbf{R}_1(-I)\mathbf{R}_3(-\omega)\mathbf{r}_o \tag{2.7}$$

Inertial to Hill (Sünkel, 1998)

$$\mathbf{r}_h = \mathbf{R}_3(u)\mathbf{R}_1(I)\mathbf{R}_3(\Omega)\mathbf{r}_i - \mathbf{r}_S \tag{2.8}$$

Local Geodetic to Geodetic (Schwarz, 1999)

$$\mathbf{r}_g = \mathbf{R}_3(\pi - \lambda)\mathbf{R}_2(\pi/2 - \varphi)\mathbf{P}_2\mathbf{r}_l + \mathbf{r}_P \tag{2.9}$$

Inertial to Geodetic (Schaub & Junkins, 2003)

$$\mathbf{r}_g = \mathbf{R}_3(\text{GAST})\mathbf{r}_i \tag{2.10}$$

where I, Ω, ω and u are Keplerian elements that respectively represent the orbit inclination, right ascension of the ascending node, perigee angle, and argument of latitude. The $\mathbf{r}_P$ vector in Equation (2.9) is the translation vector, which corresponds to the geodetic coordinates of the reference point P. As for the variable GAST shown in Equation (2.10), it represents the Greenwich Apparent Sidereal Time. Note that the equation does not include the effect of precession, nutation and polar motion. The reason is that the ephemeris parameters (Section 4.2.1) provided by the Control Segment, already accounts for these phenomena (Misra & Enge, 2001).

2.2 The Two-Body Problem

Before studying the actual satellite motion, the simplest case possible, elliptic motion, is first reviewed. In such a scenario, both the Earth and satellites are assumed to be spheres with uniform densities and all perturbation forces are neglected. Under these assumptions, the motion of a satellite revolving around the Earth obeys Kepler's law of planetary motion and Newton's law of universal gravitation (Seeber, 2003). This is known as the classical two-body central force field problem. For convenience, the subscript i is dropped in this section.

2.2.1 Kepler's Law of Planetary Motion

According to Kepler's three laws of planetary motion (Schaub & Junkins, 2003):

1. The satellite orbits are elliptical with the Earth located at one of the foci.

2. The areas swept out by the line between the satellite and the Earth are equal when the lengths of time are equal.

3. The cubes of the mean radii are proportional to the squares of the revolution periods.

The first law is relatively straightforward, so no further discussion is necessary. The second law states that satellite motion is not constant. For elliptical orbits, the further away the satellites are from the Earth, the slower they travel. The third law states that regardless of the shape of the orbits, satellites with common mean radii have the same orbital periods. In the sequel, the importance of Kepler's third law

is illustrated in the context of formation flying, as it is the key to constraining the relative motion of satellites.

2.2.2 Newton's Equation of Motion

According to Newton's mechanics, satellite orbits are defined by a homogenous differential equation of second order

$$\ddot{\mathbf{r}} + \frac{GM}{r^3}\mathbf{r} = 0, \tag{2.11}$$

where GM is the Earth's gravitational constant multiplied by its mass, $\mathbf{r}$ is the position vector of the satellites in an inertial frame and $r = \|\mathbf{r}\|$ is the distance from the centre of the Earth to the satellite. From Equation (2.11), it is clear that the acceleration and position vectors are collinear. Double integration of the equation introduces six integration constants that correspond to the initial states of the satellite. With initial conditions given, the differential equation can be integrated in time to generate the satellite's orbit at any time of interest.

2.2.3 Keplerian Elements

Instead of using position and velocity vectors, another set of orbital parameters, also known as Keplerian elements, can be used to define the elliptical motion of satellites. The Keplerian elements are much better for representing the size, shape and orientation of elliptical orbits in space, therefore it is more convenient to use them when designing the orbits with specific characteristics.

The parameter that determines the size of the ellipse is the semi-major axis a, whereas the eccentricity, e, defines the shape of the orbit. The orientation of the

ellipse in the inertial frame is defined by the inclination I which represents the angle between the orbital plane and the equatorial plane. The Right Ascension of the Ascending Node (RAAN) Ω defines the angle between the vernal equinox and the ascending node on the equatorial plane, and the argument of perigee ω is the angle from the ascending node to the perigee on the orbital plane. Finally the location of the satellite along the ellipse from the perigee is defined by the true anomaly ν. The argument of perigee and the true anomaly are sometimes combined to form a parameter called the argument of latitude u, which measures the angle of the satellite from the ascending node. This is especially useful for circular orbits, where ω and ν become inseparable (Misra & Enge, 2001). A visualization of the Keplerian elements present in the inertial coordinate system in shown in Figure 2.3

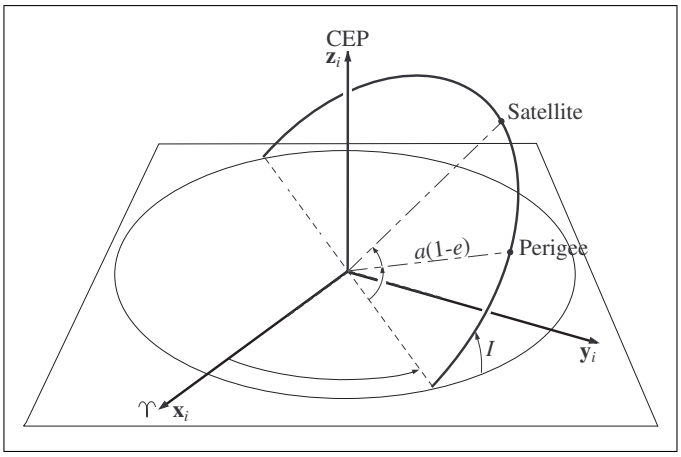


Figure 2.3: *Six Keplerian elements presented in the inertial frame (Seeber, 2003).*

2.2.4 Satellite Position and Velocity

For elliptical motion, the advantage of using Keplerian elements is that all elements, except the anomaly angle, are constant over time. Homogenous orbits can be easily

generated by propagating the mean anomaly M over time. The variation of the mean anomaly is related to the orbital frequency or mean motion of the satellite.

$$n = \sqrt{\frac{GM}{a^3}} \tag{2.12}$$

The mean anomaly at time t is given by

$$M = n(t - t_0) \tag{2.13}$$

where t_0 is the initial time at the perigee. Another anomaly measure that can be used to relate the true and mean anomalies is the eccentric anomaly, E. The mean and eccentric anomalies are related by Kepler's equation (Bate et al., 1971)

$$M = E - e \sin E \tag{2.14}$$

and the true and eccentric anomalies are related by

$$\nu = \arctan\left(\frac{\sqrt{1 - e^2}\sin E}{\cos E - e}\right). \tag{2.15}$$

The satellite position in the orbital frame is given by

$$\mathbf{r}_o = \frac{a(1 - e^2)}{1 + e\cos\nu}\begin{bmatrix} \cos\nu \\ \sin\nu \\ 0 \end{bmatrix}. \tag{2.16}$$

Equation (2.16) shows that there is no out-of-plane motion in the orbital frame. This is true for undisturbed orbits as the satellite motion is bounded by the Kepler ellipse. Taking the time derivative of Equation (2.16), the satellite velocity in the orbital frame is

$$\dot{\mathbf{r}}_o = \frac{na}{1 - e\cos E}\begin{bmatrix} -\sin E \\ \sqrt{1 - e^2}\cos E \\ 0 \end{bmatrix}. \tag{2.17}$$

The position and velocity vectors are expressed in the orbital coordinate system and they can be transformed to other systems using the coordinate transformation specified in Section 2.1.1.

2.3 J_2 Disturbed Orbits

2.3.1 Newton's Equation of Motion with J_2 Input

Up to this point, the satellite have been assumed to be in Kepler orbits, but in reality the situation is much more complicated. In order to estimate the effect of the Earth's gravity anomalies, the accelerations caused by the zonal harmonics are often considered. The disturbing potential due to the zonal harmonic is (Kaula, 2000)

$$T = -\frac{GM}{r}\sum_{n=2}^{\infty} J_n \left(\frac{a_e}{r}\right)^n P_n(\cos\theta) \tag{2.18}$$

where J_n is the zonal spherical harmonic coefficient, a_e is the equatorial radius of the Earth, P_n is the associated Legendre polynomial or function and θ is the geocentric latitude. The primary disturbance is dominated by the second zonal harmonic J_2, which resembles the equatorial bulge of the Earth. It exceeds all other zonal terms by several orders of magnitude. Therefore, restricting the expansion of the potential to $n = 2$ and taking the derivatives in the inertial frame, the accelerations due to the central gravity and J_2 are (Seeber, 2003)

$$\ddot{x} = -\frac{GMx}{r^3}\left(1 + J_2\frac{3}{2}\left(\frac{a_e}{r}\right)^2\left(1 - 5\frac{z^2}{r^2}\right)\right) \tag{2.19a}$$

$$\ddot{y} = \frac{y}{x} \tag{2.19b}$$

$$\ddot{z} = -\frac{GMz}{r^3}\left(1 + J_2\frac{3}{2}\left(\frac{a_e}{r}\right)^2\left(3 - 5\frac{z^2}{r^2}\right)\right). \tag{2.19c}$$

Using the simplicity of the Cowell's method along with numerical integration algorithms such as the Runge-Kutta or Adams-Bashforth-Adams-Moulton predictor corrector, Equation (2.19) can be numerically integrated for the determination of J_2 disturbed orbits. A detailed treatment of numerical integration solutions can be found in Montenbruck & Gill (2000).

2.3.2 Lagrange Planetary Equations of Motion

In this section the Lagrange Planetary Equations (LPE) are introduced to analyze the effect of J_2 perturbation on Keplerian elements. The LPE is a set of differential equations that express the time variations of the orbital elements as linear functions of the derivatives of the disturbing potential with respect the orbit elements. Though not shown here, the exact form of the LPE can be found in Seeber (2003). When the derivatives of the primary potential are evaluated into the LPE, the time variation of the Keplerian elements are (Sünkel, 1998)

$$\dot{a} = 0 \tag{2.20a}$$

$$\dot{e} = 0 \tag{2.20b}$$

$$\dot{I} = 0 \tag{2.20c}$$

$$\dot{\Omega} = \frac{-3nJ_2a_e^2}{2(1-e^2)^2a^2}\cos I \tag{2.20d}$$

$$\dot{\omega} = \frac{-3nJ_2a_e^2}{4(1-e^2)^2a^2}(1-5\cos^2 I) \tag{2.20e}$$

$$\dot{M} = n + \frac{3nJ_2a_e^2}{4(1-e^2)^{3/2}a^2}(3\cos^2 I - 1). \tag{2.20f}$$

The equations indicate that secular growths are functions of the orbital elements a, e and I, but ironically these parameters are not affected by the primary perturbation.

Secular drifts appear only in the orbital elements Ω, ω and M, and since their rates are constant, the precession can be determined analytically. Note that only secular effects appear in Equation (2.20). The reason is because the analysis considers only the primary second zonal harmonics, therefore minor perturbation effects are not captured by the equation.

Chapter 3

Hill Equations of Motion

Formation flying is certainly a key technology for future satellite missions, and it requires an understanding of the relative satellite dynamics for precise navigation and control (Bauer et al., 1999). In this chapter, the linearized Hill Equations (HE) are derived and utilized to study the absolute and relative motion of satellites in both homogeneous, and J_2 perturbed, gravitational fields. Furthermore, initial conditions for designing drift-free formations are developed to maintain a configuration even when exposed to J_2 perturbation. This is essential for formation flying because any insignificant drifts between the spacecraft would consume additional fuel for configuration control, which can dramatically reduce the mission lifetime (Vadali et al., 1999).

3.1 Detailed Derivation

3.1.1 Linearized Equations of Motion

The linearized HE were first presented by Hill (1878) to analyze satellite motion within the Hill frame. The derivation in Sneeuw (2002a) is followed, starting with the coordinate transformation between the inertial and Hill frames. For convenience, the subscript h is dropped in this section.

$$\mathbf{r} = \mathbf{R}\mathbf{r}_i - r\mathbf{x}, \tag{3.1}$$

where $\mathbf{R}$ is the orthogonal transformation matrix from the inertial to the Hill frame, (see Equation (2.8)) and r is the constant radius of the Hill frame. The second term is included here to consider the translation between the coordinate systems. Time derivatives can now be applied to Equation (3.1) using the product rule.

$$\mathbf{r}_i = \mathbf{R}^\mathsf{T}\mathbf{r} + r\mathbf{R}^\mathsf{T}\mathbf{x} \tag{3.2a}$$

$$\Downarrow \text{ time derivative}$$

$$\dot{\mathbf{r}}_i = \mathbf{R}^\mathsf{T}\dot{\mathbf{r}} + \dot{\mathbf{R}}^\mathsf{T}\mathbf{r} + r\dot{\mathbf{R}}^\mathsf{T}\mathbf{x} \tag{3.2b}$$

$$\Downarrow \text{ multiply by } \mathbf{R}$$

$$\mathbf{R}\dot{\mathbf{r}}_i = \dot{\mathbf{r}} + \mathbf{R}\dot{\mathbf{R}}^\mathsf{T}(\mathbf{r} + r\mathbf{x}) \tag{3.2c}$$

$$\Downarrow \text{ replace } \mathbf{R}\dot{\mathbf{R}}^\mathsf{T} \text{ by } \boldsymbol{\Omega}$$

$$\mathbf{R}\dot{\mathbf{r}}_i = \dot{\mathbf{r}} + \boldsymbol{\Omega}(\mathbf{r} + r\mathbf{x}) \tag{3.2d}$$

$$\Downarrow \text{ multiply by } \mathbf{R}^\mathsf{T}$$

$$\dot{\mathbf{r}}_i = \mathbf{R}^\mathsf{T}\dot{\mathbf{r}} + \mathbf{R}^\mathsf{T}\boldsymbol{\Omega}(\mathbf{r} + r\mathbf{x}) \tag{3.2e}$$

$$\Downarrow \text{ time derivative}$$

$$\ddot{\mathbf{r}}_i = \mathbf{R}^\mathsf{T}\ddot{\mathbf{r}} + 2\dot{\mathbf{R}}^\mathsf{T}\dot{\mathbf{r}} + \dot{\mathbf{R}}^\mathsf{T}\boldsymbol{\Omega}(\mathbf{r} + r\mathbf{x}) + \mathbf{R}^\mathsf{T}\dot{\boldsymbol{\Omega}}(\mathbf{r} + r\mathbf{x}) \tag{3.2f}$$

$$\Downarrow \text{ multiply by } \mathbf{R} \text{ and again replace } \mathbf{R}\dot{\mathbf{R}}^\mathsf{T} \text{ by } \boldsymbol{\Omega}$$

$$\mathbf{R}\ddot{\mathbf{r}}_i = \ddot{\mathbf{r}} + 2\boldsymbol{\Omega}\dot{\mathbf{r}} + \boldsymbol{\Omega}\boldsymbol{\Omega}(\mathbf{r} + r\mathbf{x}) + \dot{\boldsymbol{\Omega}}(\mathbf{r} + r\mathbf{x}). \tag{3.2g}$$

The matrix $\boldsymbol{\Omega} = \mathbf{R}\dot{\mathbf{R}}^\mathsf{T}$ is known as the Cartan matrix as it is named after the great French mathematician Cartan (1894). The rotational rate of the local orbiting triad

is described within the following matrix.

$$\mathbf{\Omega} = \begin{bmatrix} 0 & -n & 0 \\ n & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.3}$$

where n is the mean motion of the reference orbit.

Velocity in the Hill frame can be obtained by rearranging Equation (3.2d)

$$\dot{\mathbf{r}} = \mathbf{R}\dot{\mathbf{r}}_i - nr\mathbf{y} - \mathbf{\Omega}\mathbf{r}, \tag{3.4}$$

where the first term re-orients the velocity from the inertial to the Hill frame, the second term deducts the along-track motion of the reference orbit from the satellite and the last term is due to the rotating motion of the triad.

Equation (3.2g) is the second order differential equation of motion for analyzing the motion of satellites within the Hill frame. The equation shows that the acceleration in the Hill frame includes reorientation of the system's input force (dynamics) and three other inertial accelerations (kinematics). The first inertial acceleration, -2$\mathbf{\Omega}\dot{\mathbf{r}}$, is known as the Coriolis force. It is induced by motion within the rotating frame. To obtain a further understanding of the behaviour of the Coriolis force, Equation (3.3) can be substituted into the term.

$$2\mathbf{\Omega}\dot{\mathbf{r}} = 2n \begin{bmatrix} -\dot{y} \\ \dot{x} \\ 0 \end{bmatrix} \tag{3.5}$$

The Coriolis acceleration is orthogonal to the satellite motion, which means that any along-track motion can induce acceleration to the radial direction and vice-versa.

The second inertial acceleration is the centrifugal acceleration.

$$\mathbf{\Omega\Omega}(\mathbf{r} + r\mathbf{x}) = -n^2 \begin{bmatrix} x + r \\ y \\ 0 \end{bmatrix} \tag{3.6}$$

Such accelerations are directly proportional to the position of the satellites within the Hill frame. Any radial and along-track displacement can induce forces to the corresponding directions. The last inertial acceleration is known as the Euler acceleration, which is due to the inconsistent rotational rate of the triad,

$$\dot{\mathbf{\Omega}}(\mathbf{r} + r\mathbf{x}) = \mathbf{0}. \tag{3.7}$$

Since the Hill frame is assumed to be rotating at a constant speed, the Euler acceleration is therefore neglected.

The term that requires extra attention is with respect to the dynamics within the system, $\mathbf{R}\ddot{\mathbf{r}}_i$, which represents the central gravitational pull and all other perturbations that are experienced by the satellites but re-oriented to the Hill frame. To formulate the final equation of motion with simplicity, the central gravitational force is linearized at the origin of the Hill frame using a first order Taylor series expansion (Rummel, 1986)

$$\mathbf{R}\ddot{\mathbf{r}}_i = -\frac{GM}{r^3} \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} + \frac{GM}{r^3} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \mathbf{r} + \mathbf{a}, \tag{3.8}$$

where the first term is the central gravitational pull exerted on the reference orbit, the second term includes the gravitational tensor for linearization of the central field

from the origin of the Hill frame to the satellites and the last term represents all other perturbation forces acting on the satellites. Substituting Equations (3.5) to (3.8) into the differential Equation (3.2g), the final equation of motion becomes

$$\ddot{x} - 2n\dot{y} - 3n^2 x = a_x \tag{3.9a}$$

$$\ddot{y} + 2n\dot{x} = a_y \tag{3.9b}$$

$$\ddot{z} + n^2 z = a_z, \tag{3.9c}$$

which are the linearized Hill Equations of motion (Hill, 1878), also known as the Clohessy-Wiltshire equations (Clohessy & Wiltshire, 1960). As shown in Equation (3.9), the form of the HE is remarkably simple and independent of any absolute variables. Furthermore, the equations are also time-invariant and linear. This property is important because the closed-form solution of linear time-invariant systems is easier to solve if the system inputs can be expressed as periodic functions. However, due to the linearization of the central gravitational field, the HE are only valid when the reference orbits are circular and satellites are relatively close to the origin. When satellites are too far away from the Hill frame, a first order approximation of the central field will fail to model the satellite motion precisely.

Note that although the derivations here are done with absolute parameters, the HE are also applicable to study the relative motion of satellites. This can be done by simply replacing the absolute parameters with relative parameters. This procedure is feasible because of the linearity of the HE.

3.1.2 Homogenous Solution

Firstly, consider the simplest case of Kepler's motion where all perturbations acting on the satellites are neglected. Such a solution set is the homogeneous solution. To obtain the closed-form solution, methods such as the Laplace transformation, matrix exponent or induction may be applied. Note that Equation (3.9c) is decoupled, therefore its solution can be solved independently from the other two equations. The final homogeneous solution is

$$\begin{aligned} x(t) &= \frac{\dot{x}_0}{n}\sin nt - \left(3x_0 + \frac{2\dot{y}_0}{n}\right)\cos nt + \\ &\quad \frac{2}{n}(2nx_0 + \dot{y}_0) && \text{(3.10a)} \\ y(t) &= 2\frac{\dot{x}_0}{n}\cos nt + 2\left(3x_0 + \frac{2\dot{y}_0}{n}\right)\sin nt + \\ &\quad y_0 - \frac{2\dot{x}_0}{n} - 3(2nx_0 + \dot{y}_0)t && \text{(3.10b)} \\ z(t) &= \frac{\dot{z}_0}{n}\sin nt + z_0\cos nt, && \text{(3.10c)} \end{aligned}$$

where x_0, y_0, z_0, $\dot{x}_0$, $\dot{y}_0$ and $\dot{z}_0$ are initial conditions at time t_0. The fact that the cross-track component is uncorrelated makes its own solution simple; it contains only periodic terms with amplitudes dependent on its initial position and velocity. On the other hand, the general solutions in the radial and along-track directions are slightly more complicated and are tightly coupled. Not only do they contain periodic terms, but they also contain offsets and a secular drift in the along-track that can grow infinitely large as time approaches infinity. As mentioned in the previous section, the HE are valid only when the satellites are relatively close to the origin of the Hill frame. Allowing the secular drift to grow in the along-track direction causes the first order approximation to fail. Thus, the secular growth must be eliminated, resulting

in a bounded orbit constraint given in terms of Hill frame coordinates

$$2nx_0 + \dot{y}_0 = 0. \tag{3.11}$$

Furthermore, enforcing this bounded orbit constraint can also eliminate the radial offsets. Another interesting property of the general solution is that the amplitude of the periodic motion in the along-track direction is twice the amplitude of that in the radial direction. The importance of this property for configuration des'ign is demonstrated in Section 3.2.

3.1.3 Heterogenous Solution with J_2 Input

Various non-homogeneous solutions have been developed for the HE and such solutions can be found in Rummel (1986). However, the existing solutions cannot be used to represent the J_2 perturbation. To address the problem, the heterogeneous solution with J_2 input is introduced in this section. The J_2 disturbing potential can be expressed as a function of Keplerian elements by evaluating Equation (2.18) for $n = 2$ and substituting $\sin\theta$ with $\sin I \sin u$.

$$T_{20} = -\frac{1}{2} GMJ_2 \frac{a_e^2}{r^3} (3\sin^2 I \sin^2 u - 1) \tag{3.12}$$

The force functions acting on the satellites are to be approximated by the perturbation at the origin of the Hill frame. Such a zero order approximation makes the orbit elements, r and I, constant throughout, thus allowing the force functions to be expressed by sums of periodic terms. To obtain the perturbing forces in the Hill frame, the proper gradient function must be applied to the disturbing potential. For accelerations in the radial, along-track and cross-track directions, the partial derivatives

are (Sneeuw, 2000)

$$\nabla = \left[\begin{array}{ccc} \dfrac{\partial}{\partial r} & \dfrac{1}{r}\dfrac{\partial}{\partial u} & \dfrac{1}{r\sin u}\dfrac{\partial}{\partial I} \end{array} \right]^{\mathsf{T}}. \tag{3.13}$$

Applying Equation (3.13) to Equation (3.12), the J_2 perturbing forces in terms of the Hill coordinate system are

$$a_x = \frac{3}{2}k_1 \cos 2u + k_2 \tag{3.14a}$$

$$a_y = k_1 \sin 2u \tag{3.14b}$$

$$a_z = k_3 \sin u, \tag{3.14c}$$

where

$$k_1 = -\frac{3}{2}GMJ_2\frac{a_e^2}{r^4}\sin^2 I \tag{3.15a}$$

$$k_2 = -\frac{3}{2}GMJ_2\frac{a_e^2}{r^4}\left(1 - \frac{3}{2}\sin^2 I\right) \tag{3.15b}$$

$$k_3 = -\frac{3}{2}GMJ_2\frac{a_e^2}{r^4}\sin 2I, \tag{3.15c}$$

are constants. Equation (3.14) shows that both the radial and along-track forces fluctuate at twice the orbital frequency. Furthermore, the radial component consists of a constant force term. As for the cross-track, it contains only a periodic term at the orbital frequency. The HE of motion with such particular input forces can be expressed as sums of periodic functions

$$\ddot{x} - 2n\dot{y} - 3n^2 x = A_x \cos 2nt + B_x \sin 2nt + C_x \tag{3.16a}$$

$$\ddot{y} + 2n\dot{x} = A_y \cos 2nt + B_y \sin 2nt \tag{3.16b}$$

$$\ddot{z} + n^2 z = A_z \cos nt + B_z \sin nt. \tag{3.16c}$$

The constants on the right hand side can be determined analytically by defining the true anomaly of the reference orbit at the initial time. The simplest case is to start

the Hill frame at the ascending node, force several constants to become zeroes and thus simplify the expression of the final solution. Solving the equations of motion, the general solutions with J_2 input are (Xu et al., 2004)

$$\begin{aligned} x(t) &= \frac{\dot{x}_0}{n}\sin nt - \left(3x_0 + \frac{2\dot{y}_0}{n} + \frac{k_2 + \frac{5}{6}k_1}{n^2}\right)\cos nt - \frac{k_1}{6n^2}\cos 2nt + \\ &\quad \frac{2}{n}\left(2nx_0 + \dot{y}_0 + \frac{k_1 + k_2}{n}\right) \qquad (3.17a) \\ y(t) &= 2\frac{\dot{x}_0}{n}\cos nt + 2\left(3x_0 + \frac{2\dot{y}_0}{n} + \frac{k_2 + \frac{5}{6}k_1}{n^2}\right)\sin nt - \frac{k_1}{12n^2}\sin 2nt + \\ &\quad y_0 - 2\frac{\dot{x}_0}{n} - 3\left(2nx_0 + \dot{y}_0 + \frac{\frac{1}{2}k_1 + \frac{1}{3}k_2}{n}\right)t \qquad (3.17b) \\ z(t) &= \left(z_0 - \frac{k_3}{2n}t\right)\cos nt + \left(\frac{\dot{z}_0}{n} + \frac{k_3}{2n^2}\right)\sin nt. \qquad (3.17c) \end{aligned}$$

Comparing the general solution to the homogeneous solution, Equation (3.10), the amplitudes at the orbital frequency in the along-track are still twice the amplitudes in the radial direction but the magnitudes are slightly deviated. Additional periodic motions at twice the orbital frequency are introduced to the radial and along-track directions. As for the drift in the along-track, it no longer satisfies the bounded orbit constraint given in Equation (3.11) and is not equal to the radial offset. The drift term here is not the bounded orbit constraint for the J_2 gravitational field. Forcing the drift to zero only matches the frequency of the satellites with the Hill frame. In the cross-track direction, only the periodic terms exist, but once again the amplitudes are perturbed. An input force at the orbital frequency in the cross-track results in resonance, which is represented by the drift term within the amplitude.

Note that Equation (3.17) is only applicable to the analysis of absolute orbits. When applied to relative motion, the solution simply reduces to the homogeneous solutions in Equation (3.10) because the relative J_2 perturbation is zero due to the

zero order approximation.

3.2 Configuration Design

In this section, the use of the HE for formation design is demonstrated. Two relative configurations are introduced and their immunity towards the J_2 perturbation is investigated.

3.2.1 Design Parameters and Orbit Element Difference

Using the Hill coordinates for formation design is problematic because position and velocity vectors are only instantaneous and do not describe the satellite motion over time. To address this problem, the homogeneous solution of the HE can be expressed in the following form (Vadali et al., 2000).

$$x_h(t) = A\cos(nt+\alpha) + x_{\text{off}} \tag{3.18a}$$

$$y_h(t) = -2A\sin(nt+\alpha) + y_{\text{off}} - \frac{3}{2}ntx_{\text{off}} \tag{3.18b}$$

$$z_h(t) = B\cos(nt+\beta) \tag{3.18c}$$

where A and B are the in-plane and out-of-plane amplitudes, α and β are the corresponding phase angles, and x_{off} and y_{off} are the radial and the along-track offsets. Introducing these design parameters brings great convenience to formation design because the configurations are much easier to visualize when represented with amplitude and phase than with position and velocity vectors at a single location. The relationship between the initial Hill coordinates and the design parameters can be analytically determined from Equations (3.10) and (3.18).

According to Schaub (2002), if the reference orbit is assumed to be circular and the Hill coordinates are relatively small compared to the orbit radius, then a direct relationship between the relative orbit elements and the design parameters can be established. Here, the reverse transformation from the design parameters to the orbit element differences is presented.

$$\Delta a = x_{\text{off}} \tag{3.19a}$$

$$\Delta e = \frac{A}{a} \tag{3.19b}$$

$$\Delta I = \frac{\mp B \tan(\omega - \beta)}{a\sqrt{\tan^2(\omega - \beta) + 1}} \tag{3.19c}$$

$$\Delta \Omega = \frac{\pm B}{a \sin I \sqrt{\tan^2(\omega - \beta) + 1}} \tag{3.19d}$$

$$\Delta \omega = \pi - \alpha \tag{3.19e}$$

$$\Delta M_0 = \frac{y_{\text{off}}}{a} - \Delta\omega - \Delta\Omega \cos I. \tag{3.19f}$$

Note that the design parameter x_{off} is the radial offset which is also associated with the along-track drift. Therefore, the bounded orbit constraint in terms of orbit element differences is $\Delta a = 0$. This makes sense because orbits with identical semi-major axes have identical periods. Any discrepancies in their orbital periods introduce secular growth to the relative motion. Equation (3.19) shows that the size and phase of the in-plane motion are associated with the elements Δe and $\Delta\omega$ correspondingly. In the case of the out-of-plane motion, it is controlled by both ΔI and $\Delta\Omega$. The last element ΔM_0 deals with the offset in the along-track direction and acts as the responding parameter to the other elements to fulfill the requirements necessary for configuration design.

3.2.2 2-by-1 Elliptical Formation

Analyzing Equation (3.18), it is clear that the periodic motion in the along-track is orthogonal to the radial with twice the amplitude. When all drifts and cross-track motions, are eliminated by letting $x_{\text{off}} = B = 0$, the result is an in-plane 2-by-1 elliptical relative motion, also known as the interferometric cartwheel (Massonnet, 2001). To minimize the distance between the satellites and the Hill frame, and to allow the relative motion to revolve around the origin, the along-track offset must be set to zero, $y_{\text{off}} = 0$. Substituting the conditions into Equation (3.19), the non-zero orbital element differences are

$$\Delta e = \frac{A}{a} \tag{3.20a}$$

$$\Delta\omega = \pi - \alpha \tag{3.20b}$$

$$\Delta M_0 = \alpha - \pi. \tag{3.20c}$$

With a constant amplitude A defined, different satellites can be placed in the 2-by-1 ellipse by varying the phase angle α, Figure 3.1. Satellites in such formations

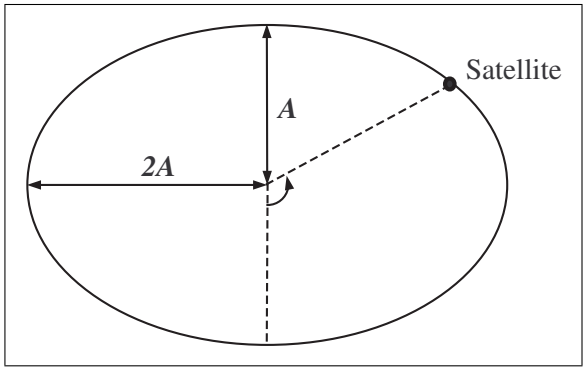


Figure 3.1: *Key design parameters of 2-by-1 elliptical formation.*

have a common eccentricity and the variations are only in the perigee and the mean

anomaly.

Recall that the secular growth of the orbital elements due to the primary J_2 perturbation is dependent on the semi-major axis, eccentricity and inclination. For satellites in the 2-by-1 elliptical formation, these three elements are identical, which means that their secular drifts due to J_2 are equal. This property is especially important since it confirms that 2-by-1 elliptical formations are actually J_2 invariant relative orbits.

3.2.3 Circular Formation

The other configuration to be considered is the circular relative motion. Satellites in such formations revolve in a circular motion around the origin of the Hill frame and the plane of motion is inclined at 30 degrees from the horizontal plane. Under homogeneous conditions, the inter-satellite distances are constant at all times. The circular motion is generated by setting the amplitude of the cross-track motion to exactly $\sqrt{3}$ of the radial amplitude, $B = \sqrt{3}A$, and letting the out-of-plane phase angle equal the in-plane phase angle, $\beta = \alpha$ (Alfriend et al., 2000b). Similar to the 2-by-1 elliptical motion, the offset parameters are again set to zero $x_{\text{off}} = y_{\text{off}} = 0$. Applying the conditions to Equation (3.19), the orbit element differences of circular relative motion are

$$\Delta a = 0 \tag{3.21a}$$

$$\Delta e = \frac{A}{a} \tag{3.21b}$$

$$\Delta I = \frac{\mp\sqrt{3}A\tan(\omega-\alpha)}{a\sqrt{\tan^2(\omega-\alpha)+1}} \tag{3.21c}$$

$$\Delta\Omega = \frac{\pm\sqrt{3}A}{a\sin I\sqrt{\tan^2(\omega-\alpha)+1}} \tag{3.21d}$$

$$\Delta\omega = \pi - \alpha \tag{3.21e}$$

$$\Delta M_0 = \alpha - \pi - \Delta\Omega\cos I. \tag{3.21f}$$

The radius of the circular motion is $2A$ and α defines the phase angle along the circular path. The cross-track motion is created by the non-zero ΔI and $\Delta\Omega$; their values can vary depending on the phase angle and the perigee. Equation (3.21) indicates that satellites within the circular formation have various inclination differences. Since the secular drift due to J_2 is dependent on the inclination, the circular formation is not J_2 invariant in general.

There is one situation where the circular motion can be J_2 invariant. To make this possible the inclination differences ΔI must become zero and all the cross-track motions must be created by $\Delta\Omega$. The conditions that satisfy such requirements are $\alpha = \omega$ and $\alpha = \omega + 180$, which limits the formation to hold only two satellites. Another circumstance that needs to be considered is the precession of the perigee due to J_2, which violates the conditions necessary for circular relative motion specified in Equation (3.21). To avoid such a problem, the motion of the perigee must be dampened by setting the orbits to a critical inclination of $I = 63.43°$, but the negative consequence is a constrained satellite coverage.

3.3 Evaluation of the Hill Equations of Motion

The purpose of this section is to evaluate the accuracy of the HE of motion for modelling the absolute and relative motion of satellites in both the central and

disturbed gravitational fields. The evaluation is done using a simulated software developed with MATLAB®. When initial conditions are provided, the software can generate homogeneous orbits by propagating the anomaly over the specified time duration. As for J_2 disturbed orbits, they are simulated by numerical integration of Equation (2.19) and the integration algorithm chosen is the Runge-Kutta single step method. The software package also includes all kinds of coordinate transformation functions for analysis of satellite orbits in different reference frames.

3.3.1 Absolute Motion in Central Gravitational Field

The HE equations are linearized at the origin of the orbiting triad, therefore, the accuracy of the solutions is directly influenced by the displacement from the origin. To quantify the precision, satellites with variable distances from the origin are simulated under homogeneous conditions, see Table 3.1. The three satellites simulated are in circular motion with radii of 0 km, 6 km and 10 km, see Figure 3.2. Their

Table 3.1: *Orbit elements of the Hill frame and the satellite design parameters.*

Orbit Elements	Hill Frame	Design Parameters	Satellite 1 (Green)	Satellite 2 (Red)	Satellite 3 (Blue)
a	$a_e + 500$ km	A	0 km	3 km	5 km
e	0	B	0 km	$3\sqrt{3}$ km	$5\sqrt{3}$ km
I	89°	α	0	0	0
Ω	0°	β	0	0	0
ω	0°	x_{off}	0 km	0 km	0 km
M	0°	y_{off}	0 km	0 km	0 km

orbits are first generated in the inertial frame by propagating the mean anomaly over a duration of six orbits. The motions, in terms of Hill coordinates, are then obtained by using the direct linear mapping developed by Alfriend & Schaub (2000).

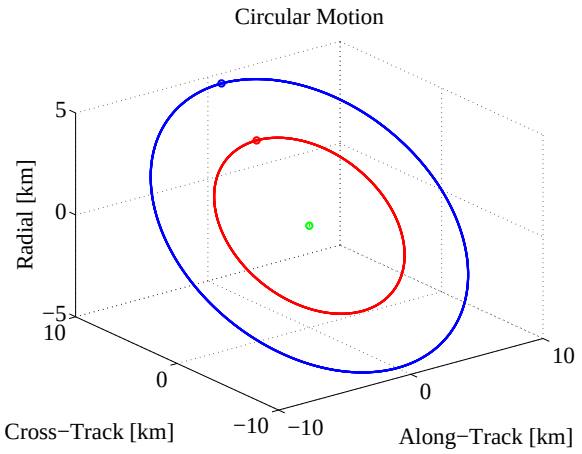


Figure 3.2: *Satellite trajectories in circular motion presented in the Hill frame.*

These trajectories act as the reference and are used to evaluate the precision of the homogeneous solution of the HE, Equation (3.10).

The position and velocity errors of the HE are shown in Figure 3.3. The only errors visible are the position estimates in the along-track. These errors are completely periodic and the levels are directly proportional to the displacement of satellites from the origin. For a circular motion with a 10 km radius (blue), the maximum along-track position error can reach 5 m, hence it is quite insignificant. The results indicate that the HE are relatively accurate for modelling the absolute motion of satellites in a central gravitational field. Error appears only in the along-track position, while all other components are modelled perfectly (zero error).

3.3.2 Relative Motion in a Central Gravitational Field

In the previous section, modelling the absolute motion of satellites using the linearized HE of motion was examined. A more critical issue is to test the equations with relative motion. Such investigations can evaluate the applicability of the HE for

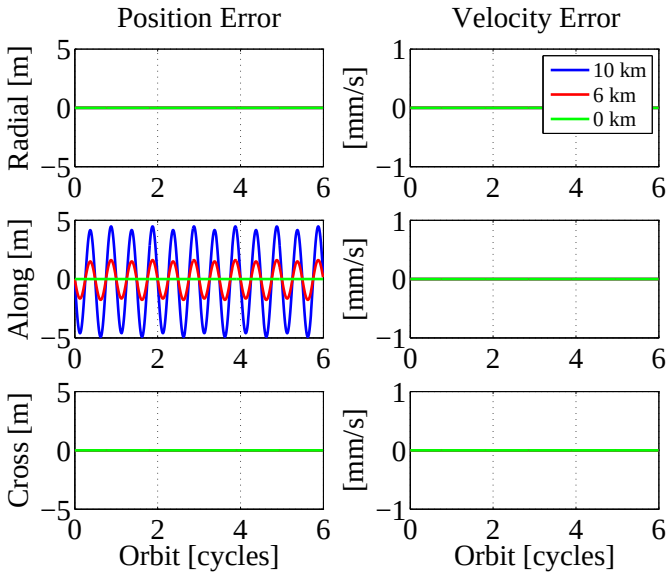


Figure 3.3: *Errors committed by the* HE *for satellites in circular motion.*

modelling the relative motion of satellites under the central gravitational field. Using the same Hill frame specified in Table 3.1, the design parameters of the satellites are shown in Table 3.2. The configuration is a circular formation consisting of two satellites with a 12 km separation, see Figure 3.4.

Table 3.2: *Satellite design parameters in circular formation.*

Design Parameters	Satellite 1 (Red)	Satellite 2 (Blue)
A	3 km	3 km
B	$3\sqrt{3}$ km	$3\sqrt{3}$ km
α	0	π
β	0	π
x_{off}	0 km	0 km
y_{off}	0 km	0 km

The relative position and velocity errors of the HE are shown in Figure 3.5. The only errors made by the HE are in the relative positions in the along-track, which

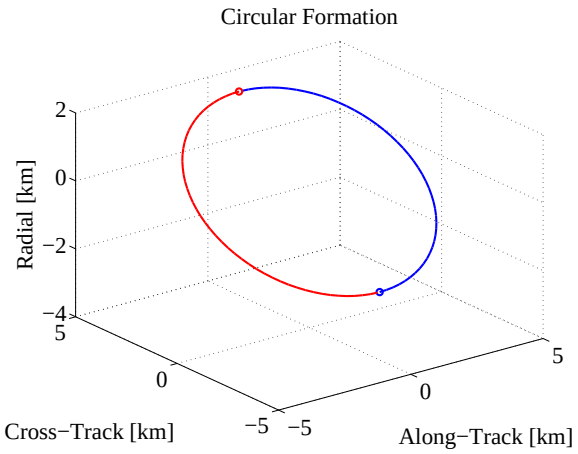


Figure 3.4: *Satellite trajectories in circular formation presented in the Hill frame.*

is similar to the previous analysis with absolute motion. For the circular formation with a 12 km separation, the relative position errors in the along-track direction are totally periodic and the maximum is 0.5 mm, which is insignificant. No errors are committed in the other directions. These results indicate that the HE are extremely accurate for modelling the relative motion of satellites in the central gravitational field.

3.3.3 Absolute Motion in J_2 Disturbing Field

In this section, the solutions of the HE with J_2 input are evaluated. Such investigations can confirm the accuracy of the equations of motion for modelling J_2 disturbed orbits. To allow some precession to the ascending node, the inclination of the Hill frame is set to 85°. The initial Hill coordinates of the satellites, except the along-track velocity, are all zero, see Table 3.3. The along-track velocity is specifically chosen in order to eliminated the along-track drift in Equation (3.17). The motion of the perturbed satellite orbits are shown in Figure 3.6. To generate the disturbed

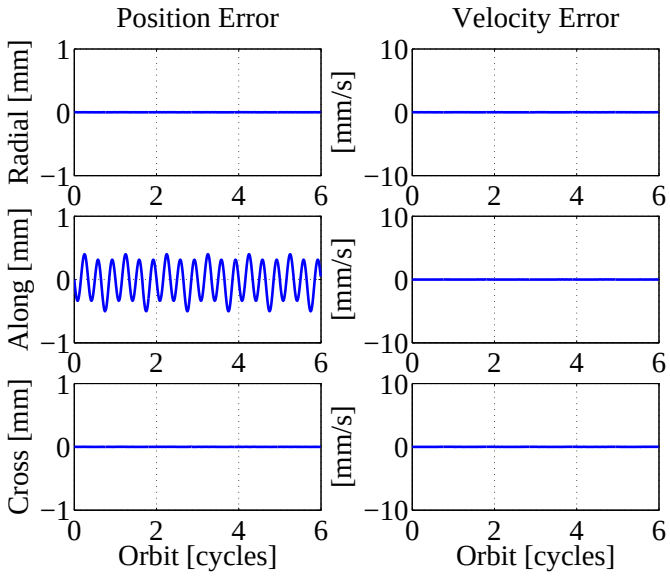


Figure 3.5: *Relative errors committed by the* HE *for satellites in circular formation.*

orbit, the Hill coordinates are first transformed back to the inertial frame using the inverse of direct linear mapping. With the initial conditions, Equation (2.19) are integrated numerically to produce the orbit trajectory. The resulting orbit is then transformed back to the Hill frame. The motions in the radial and along-track directions are periodic without any signs of drift. Resonance in the cross-track motion is definitely visible and this is the direct cause of the precession of the ascending node due to J_2. The increasing amplitude is actually bounded as the ascending node would precess back to its original state. However, the bounding limit is far too large and becomes unapplicable for interpretation in the Hill frame.

The position and velocity errors are shown in Figure 3.7. Although the drift term in the along-track is already considered, the satellite is still slowly drifting away from the Hill frame. After only two revolutions, the along-track drift is greater than 100 m. The amplitude of the cross-track error is growing with time. The

Table 3.3: *Orbital elements of the Hill frame and the satellite initial Hill coordinates.*

Orbit Elements	Hill Frame	Hill Coordinates	Satellite (Blue)
a	$a_e + 500$ km	x_0	0 km
e	0	y_0	0 km
I	85°	z_0	0 km
Ω	0°	$\dot{x}_0$	0 m/s
ω	0°	$\dot{y}_0$	1.812 m/s
M	0°	$\dot{z}_0$	0 m/s

radial error might seem periodic, but if the number of revolutions is extended, it is expected to grow over time as well, because the satellite drift causes the zero order approximation to be invalid. The developed equations of motion do not capture the entire J_2 perturbation, therefore the model works well for only a short duration ($<$ 5 min.).

3.3.4 Relative Motion in J_2 Disturbing Field

A homogeneous solution of the HE is tested in this section for modelling the relative motion of satellites under the influence of J_2 perturbation. The configuration chosen is the 2-by-1 elliptical formation consisting of two satellites with separation distances that vary from 2 to 4 km. The design parameters are given in Table 3.4 and the configuration is illustrated in a perturbed circular orbit, see Figure 3.8. To generate the J_2 disturbed orbits, the design parameters are first transformed to orbital element differences using Equation (3.19). The resulting orbit elements represent only the mean elements. For orbit integration, the instantaneous (osculating) orbital elements are preferred (Schaub & Alfriend, 1999). With the osculation elements determined,

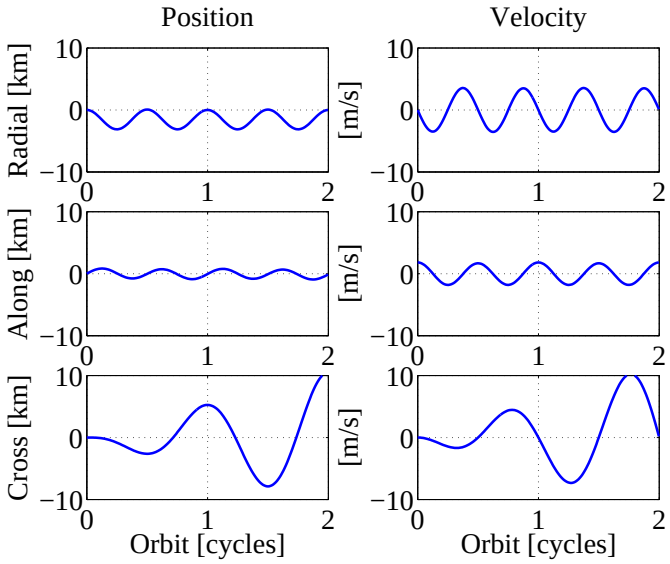


Figure 3.6: *J_2 disturbed orbits presented in the Hill frame.*

Table 3.4: *Orbit parameters of the Hill frame and the 2-by-1 elliptical formation.*

Orbit Elements	Hill Frame	Hill Coordinates	Satellite 1 (Blue)	Satellite 2 (Red)
a	$a_e + 500$ km	A	1 km	1 km
e	0	B	0 km	0 km
I	85°	α	0	π
Ω	0°	β	0	π
ω	0°	x_{off}	0 km	0 km
M	0°	y_{off}	0 km	0 km

the orbits are numerically integrated using Equation (2.19).

The relative position and velocity errors are shown in Figure 3.9. Even though the elliptical formation is J_2 invariant, a secular drift in the along-track direction is still modelled by the HE. The along-track drift reaches almost 10 m after only one tenth of a cycle and since it is coupled with the radial motion, a similar error level is found in the radial direction. As for the cross-track error, it is insignificant compared to the errors in the other directions. This is expected because the out-of-

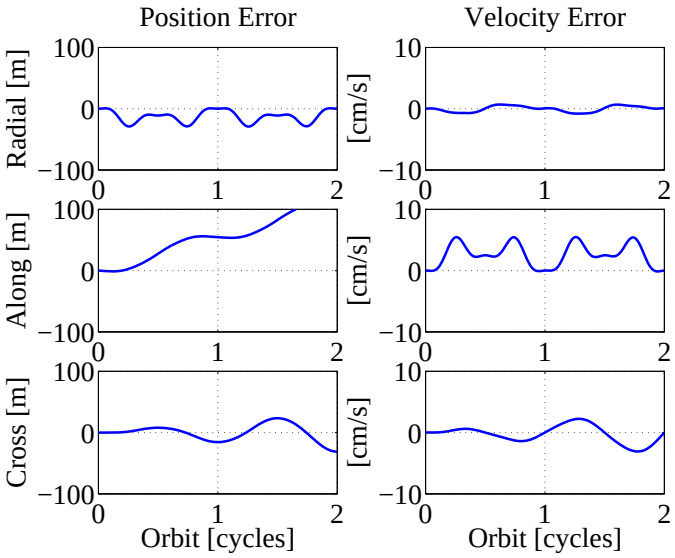


Figure 3.7: *Errors committed by the non-homogeneous* HE *for a J_2 disturbed orbit.*

plane motion is limited in the elliptical formation. In general, the results indicate that the homogeneous HE can only model J_2 disturbed relative motion over a short period of time ($< 5\,\text{min.}$).

3.4 Summary

A thorough analysis of the HE has been conducted in this chapter. Different solution forms were introduced to describe the satellite motion in both central and disturbed gravitational fields. Two formation flying designs have also been analyzed in terms practicability for future satellite missions. Some of the conclusions drawn from such analysis are:

- The 2-by-1 elliptical formation is applicable to future formation flying missions due to its J_2 invariant characteristics. However, the circular formation is not as practical since various orbital constraints are needed to prevent the configura-

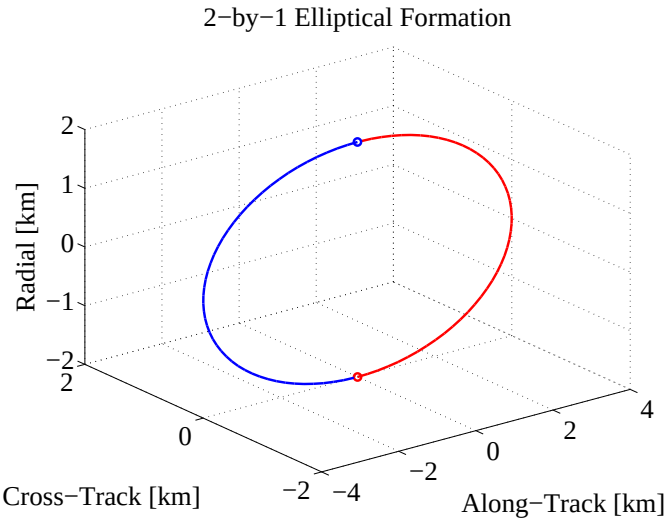


Figure 3.8: *J_2 disturbed 2-by-1 elliptical formation in a perturbed reference orbit.*

tion from drifting. The only situation that would make the circular formation feasible is to place the satellites at an extremely high orbit to minimize the perturbation effect of J_2.

- The HE are capable of modelling the absolute satellite motion in a central gravitational field without any drift. Many of the model errors are found in the along-track position and the level of error is proportional to the formation size. As for modelling the relative satellite motion in a undisturbed field, they are even more precise. In the relative case, sub-millimetre position accuracy can be achieved.

- The modelling of absolute and relative satellite motion in the J_2 disturbed gravitational field with the HE is not feasible for extended periods. The modelling accuracy degrades as time increases. Relative position errors can exceed the metre level after only one tenth of an orbit. Therefore, the use of the HE

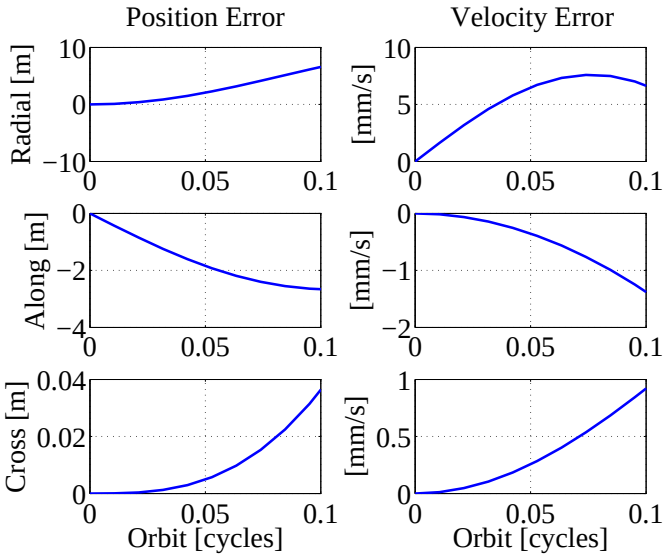


Figure 3.9: *Relative errors committed by the homogeneous* HE *for a J_2 disturbed orbit.*

should be limited to short durations ($< 5\,\text{min.}$) in a J_2 perturbed field.

Chapter 4

GPS Measurements and Error Sources

In this chapter, the basic measurements provided by GPS are discussed. Observation equations are formulated to assess how associated error sources corrupt these measurements. Furthermore, relative position models are introduced to reduce the impact of these errors.

4.1 GPS Measurements

4.1.1 Code Phase

The code phase, also known as the pseudorange, is the distance measurement generated by multiplying the apparent transit time of the GPS signal by the speed of light.

$$p = c\delta t, \tag{4.1}$$

where p is the pseudorange in metres, c is the speed of light and δt is the apparent transit time. The apparent transit time is measured by determining the time shift required to align the receiver's replica of the Pseudo Range Noise (PRN) code with the incoming signal. There are two sets of PRN codes transmitted by the GPS satellite, which are the Course/Acquisition (C/A) and Precision (P) codes (Spilker Jr., 1996). The C/A code is uniquely designed for every satellite and repeats every millisecond, which is equivalent to a chip width of 300 m. As for the P code, the chipping rate is

ten times that of the C/A code, which corresponds to a 30 m chip width. The smaller chip width of the P code provides better precision in the pseudorange measurements. However, the P code is encrypted to limit its access to authorized users.

Equation (4.1) represents only an ideal situation where the pseudorange is not contaminated by errors, which is impossible in reality. The error sources affecting the pseudorange include the satellite and receiver clocks, the satellite orbits, the atmosphere, the environment and the receiver itself. Taking these errors into consideration for the pseudorange measurement, the observation equation becomes

$$p = \rho + d\rho + d_{\text{ion}} + d_{\text{trop}} + c(dt - dT) + \varepsilon_p, \tag{4.2}$$

where $d\rho$ is the orbital error, d_{trop} and d_{ion} are the atmospheric delays associated with the troposphere and the ionosphere, dt and dT are satellite and receiver clock errors respectively, and ε_p represents all other un-modelled effects such as multipath and receiver noise. Equation (4.2) illustrates that the pseudorange is actually a biased and noisy measurement of the true range. The navigation accuracy obtained from these measurements depends on how well the errors can be modelled or eliminated.

4.1.2 Carrier Phase

The receiver is not only able to replicate the PRN code to measure the pseudorange, but it can also replicates the carrier to accurately measure the fractional difference of the phase from the received signal. This fractional phase difference, when combined with the integrated Doppler measurements (beat frequency) over time produces the carrier phase measurement, hence it is also known as the accumulated phase or the integrated Doppler. Since the fractional phase can be measured only after signal

acquisition, the initial number of whole cycles between the satellite and the receiver is unknown, thus, making the measurement ambiguous. If the carrier tracking loop maintains lock on the incoming signal, then the integer ambiguity remains constant; however, any breaks of lock will cause the value to change.

Accounting for all error sources affecting the carrier phase measurements, the observation equation in cycles is

$$\phi = [\rho + d\rho - d_{\text{ion}} + d_{\text{trop}} + c(dt - dT)]\,\lambda^{-1} + N + \varepsilon_{\phi}, \tag{4.3}$$

where λ_{ϕ} is the wavelength of the carrier signal, N_{ϕ} is the integer ambiguity and ε_{ϕ} represents multipath and receiver noise. When comparing this observation equation to Equation (4.2), it is evident that they are similar in that both are corrupted by the same error sources. Nonetheless, it is important to realize that differences between the two equations do exist and these differences are due to the ionosphere, and the integer ambiguity. Also, the multipath and receiver noise errors on the carrier phase measurement are much smaller than the code. The resolution of this ambiguity will depend on how well the errors can be reduced or modelled.

4.2 GPS Error Sources

GPS errors sources can be categorized into the following groups:

- errors in the broadcast ephemeris and clock parameters contained in the navigation message,
- uncertainties regarding the apparent transit time of the signal through the atmosphere and

- interference from reflected signals and receiver noise affecting the precision of the measurements.

In this section, the error sources corrupting GPS measurements are discussed. The objective is to understand the behaviours and the magnitudes of these errors, and also to look at different methods for error mitigation.

4.2.1 Broadcast Ephemeris

GPS navigation is accomplished through trilateration, where the position of the receiver is estimated based on distance measurements from the GPS satellites. The technique of trilateration also requires the positions of the satellite, which are expressed by an expanded set of quasi-Keplerian parameters that consists of 15 elements (ICD-GPS-200c, 1997). These parameters, together with a reference epoch, form the ephemeris parameters in the GPS navigation message, see Table 4.1. The ephemeris parameters are estimated based on a gravity model to extrapolate their values into the future. The prediction error made by the model is the orbital error. Since the error grows over time, the broadcast ephemeris is updated every two hours and new information is uploaded onto the satellites on a daily basis. The typical Root-Mean-Square (RMS) of orbital error for the broadcast ephemeris is roughly 2 m (IGS, 2005).

For relative positioning, the influence of the orbital errors on the baseline is purely geometrical and can be approximated as (Wells et al., 1986)

$$\frac{db}{b} = \frac{d\rho}{\rho}, \tag{4.4}$$

where db is the baseline error and b is the baseline distance. Equation (4.4) indicates

Table 4.1: *Description of ephemeris parameters in the* GPS *navigation message.*

parameter	description
t_{oe}	ephemeris reference of time
$\sqrt{a}$	square root of the semi-major axis
e	eccentricity
i_0	inclination angle at reference time
Ω_0	longitude of ascending node at reference time
ω	argument of perigee
M_0	mean anomaly at reference time
Δn	correction to mean motion
$\dot{i}$	rate of change of inclination angle with time
$\dot{\Omega}$	rate of change of RAAN with time
C_{uc}, C_{us}	amplitudes of harmonic correction terms for argument of latitude
C_{rc}, C_{rs}	amplitudes of harmonic correction terms for orbit radius
C_{ic}, C_{is}	amplitudes of harmonic correction terms for inclination angle

that the effect of orbital errors on baselines is limited. The true range ρ between the satellites and the receivers varies from approximately 20,000 km, when the satellite is at the zenith, to about 26,000 km, when the satellite is near the horizon. Assuming a satellite is at the zenith and has 2 m of orbital error, the corresponding baseline error is 0.1 parts-per-million (ppm). For a baseline distance of 10 km, the corresponding baseline error would be 1 mm, which is negligible.

4.2.2 Ionosphere

The atmosphere that extends from 50 to 1,000 km above the Earth's surface is known as the ionosphere. This region is filled with free electrons formed by the ionization of molecules due to the Sun's radiation (Bugoslavskaya, 1962). Therefore, the intensity of the ionosphere is directly related to the amount of solar activity. The ionosphere

is composed of several different layers, named D, E, F_1 and F_2 (Rishbeth & Garriott, 1969). These layers are separated according to differences between their altitude and electron density. The F_2 layer is approximately 450 km above Earth's surface which is where the peak electron density occurs (Liao, 2000).

The speed of an electromagnetic wave in the ionosphere is dependent upon the number of free electrons in its path. The Total Electron Content (TEC) is used to quantify the number of electrons in a vertical column extending from a satellite to a receiver with a cross-sectional area of $1\,\mathrm{m}^2$. The ionospheric delay of a code phase measurement in metres can be expressed as (Skone, 1998)

$$d_{\mathrm{ion}} = \frac{40.3 \cdot \mathrm{TEC}}{f^2}, \tag{4.5}$$

where the value of 40.3 is an empirically derived constant and f is the frequency of the carrier signal. The value of the TEC will vary depending on the hour of the day and the season of the year. The diurnal maximum is at 2:00 pm local time when the rate of ionization is highest. The seasonal maximum is during the spring equinox, when the radiation is more direct. The TEC also depends on the level of solar activity which has a cycle of approximately 11 years (Klobuchar, 1996). During the solar maximum, the level of ionization can increase by three times compared to the minimum. There is also a geographic dependence of the TEC. The TEC level at the auroral zone or geomagnetic equator is normally higher and scintillation can occur more often around these regions. Scintillation is the fluctuations in the amplitude and phase of the GPS signal due to small-scale irregularities in the electron content (Wanninger, 1993). In such situations, the signal-to-noise ratio is significantly lower than usual, which can cause cycle slips or loss of lock.

The ionospheric advance of the carrier phase measurement in metres is equal to the code phase delay in magnitude but opposite in sign. This imposes a problem for single-frequency receivers when using the carrier phase to smooth the pseudorange, since code-carrier divergence can occur if the smoothing interval is too long (Misra et al., 1999). From Equation (4.5), it is evident that the ionospheric error is frequency dependent. This means that the ionospheric delay and advance can be estimated with dual-frequency receivers, which can remove the first order ionospheric effect from the measurements. The L_1 ionospheric correction estimated using dual-frequency code phase measurements is (Liu, 2004)

$$d_{\text{ion}}^{p} = \frac{f_{\text{L2}}^2}{f_{\text{L1}}^2 - f_{\text{L2}}^2}(p_{\text{L2}} - p_{\text{L1}}). \tag{4.6}$$

The estimated correction can remove 99% of the first order ionospheric effect, however this will also amplify the uncorrelated errors by approximately three times (Misra & Enge, 2001). A more accurate estimation of the ionospheric correction can be done using the dual-frequency carrier phase measurements (Liao, 2000).

$$d_{\text{ion}}^{\phi} = \frac{f_{\text{L2}}^2}{f_{\text{L1}}^2 - f_{\text{L2}}^2}\left[\lambda_{\text{L1}}(\phi_{\text{L1}} - N_{\text{L1}}) - \lambda_{\text{L2}}(\phi_{\text{L2}} - N_{\text{L2}})\right]. \tag{4.7}$$

The correction estimated by carrier phase measurements is less noisy when compared to Equation (4.6), but it requires the integer ambiguities of both L_1 and L_2, hence, making it impractical for absolute ionospheric error estimation. However, if both L_1 and L_2 carrier tracking is maintained with no cycle slips, then the differential ionospheric delay over time can be determined because the ambiguities remain unchanged.

Klobuchar (1986) developed an ionospheric model that is currently used by the GPS control centre as part of the navigation message. Such a model is important for

single-frequency receivers where ionospheric error estimation cannot be done. The Klobuchar model provides the zenith delay with a constant value at nighttime and a half-cosine function during the daytime to represent the daily maximum. Since only the zenith delay is provided, an obliquity factor as a function of satellite elevation must be applied to compute the actual error. According to Feess & Stephens (1987), the broadcast model is estimated to reduce the RMS ionospheric effect by only 50% at mid latitude regions. The remaining error in zenith delay can reach 10 m level during the day at mid-latitudes region and could be worse during peaks of solar activities.

The level of the ionospheric error is dependent upon the TEC along the signal path. For receivers that are not far apart (< 100 km), the signal paths would be similar, because the signal is transmitted from GPS satellite 20,000 km away. Taking advantage of such a property, the ionospheric error can be reduced effectively through relative positioning. The level of the relative ionospheric error would depend upon the spatial variability of the TEC within the ionosphere. When the level of ionospheric activity is low, a typical gradient of the TEC can result in relative ionospheric errors in the order of 1-2 ppm (Parkinson, 1996). In conditions when the ionosphere is more active, errors in the order of 10 ppm can be expected and even more significant during solar magnetic storm (Foster, 2000).

4.2.3 Troposphere

The troposphere is the atmosphere up to 60 km above the Earth's surface. As GPS signals travel through the troposphere, water particles cause the signal path to refract and bend slightly. Such refractivity decreases the speed of the electromagnetic

wave, thus, causing a delay in the signal. Unlike the ionosphere, the tropospheric delay is non-dispersive at L_1 and L_2 frequencies. Instead, it is dependent upon the atmospheric temperature, pressure and relative humidity (Shrestha, 2003). The variability of the troposphere is quite low, thus, the error can be reduced to the centimetre level when models are applied. Many models are available to estimate the tropospheric delay and one is the Hopfield model (Hopfield, 1970).

For satellite formation flying, the spacecraft of interest are within the LEO altitude, which is well above the troposphere and therefore, not effected by it. Thus, the tropospheric error will be not discussed in detail herein.

4.2.4 Multipath

Multipath occurs when one or more reflected signals interfere with the direct line-of-sight signal at the antenna (Braasch & van Graas, 1991). The reflected signals are delayed and are usually weaker than the direct signals, such that the measurement error due to multipath is dependent upon this signal strength and the delay between the reflected and direct signals. They affect both the code and carrier phase measurements, but the magnitudes of their impact are significantly different. The maximum multipath code phase delay is half a code chip (Ray, 2000). For the C/A code, this is equivalent to 150 m. This means that the signal structure itself is immune to reflectors more than 150 m away for the C/A code. Typical multipath errors in code phase measurements vary from 1 m in a clean environment to more than 5 m in a highly reflective environment. The multipath errors in carrier phase measurements are typically two orders of magnitude smaller than the code. The maximum magnitude cannot exceed one quarter of the carrier wavelength (Georgiadou & Kleusberg,

1988), which is equivalent to 5 cm for L_1 and 6 cm for L_2.

Multipath is not correlated beyond a few centimetres (Ray, 2000), therefore differential techniques cannot reduce its effect; nonetheless, there are many other methods to reduce the effect of multipath. The simplest method is to carefully select the antenna site and to avoid any potential reflectors, but this is not always possible. If the site selection is limited, then users should consider GPS equipment with multipath mitigation technologies. The gain pattern of the antenna can be specially designed to reject multipath at low elevations, but the disadvantage is reduced tracking ability. Attaching a groudplane and chokering to the antenna are other options to be considered. Regarding the receivers, users can select receivers with narrow correlators to reduce the multipath errors and eliminate some of the long multipath. However, the narrow correlators are not perfect; they are still vulnerable to short multipath signals (Van Dierendonck et al., 1992).

For satellite formation flying mission where configuration size exceeds 150 m, the only significant reflective source for multipath is the spacecraft itself. Thus, the maximum path delay is dependent upon the dimension of the spacecraft. Kroes (2006) has analyzed the multipath error for spacecraft from the GRACE mission. The analysis indicates that the C/A code multipath error in spaceborne environment is at the decimetre level. As for the carrier phase, Montenbruck & Kroes (2003) have analyzed the GPS data from the CHAMP mission. The research has been conducted to assess the performance for the BlackJack receiver and found the carrier phase multipath error in the millimetre level.

4.3 Differential Observation

For GPS positioning, the highest accuracies can be obtained only after the carrier phase ambiguities are resolved to their integer values. The ability to estimate these integers is dependent upon the level of errors contaminating the measurements. In this section, a method known as differential positioning is presented to reduce or even completely eliminate the dominant error sources contaminating the measurements.

4.3.1 Single Difference

Suppose that there is a receiver with known coordinates stationed nearby the user's receiver, which is observing satellites similar to the user. Under this circumstance, Single Difference (SD) carrier phase measurements may be formed by taking the difference between the measurements of the rover (user's receiver) and the base (receiver with known coordinates).

$$\begin{aligned} \Delta\phi &= \phi_{\text{rover}} - \phi_{\text{base}} \\ &= \left[\Delta\rho + \Delta d\rho - \Delta d_{\text{ion}} + \Delta d_{\text{trop}} + c\Delta dT\right]\lambda^{-1} + \Delta N + \Delta\varepsilon_{\phi}, \end{aligned} \tag{4.8}$$

where Δ is the SD operator. By forming the SD measurements, the estimation becomes a relative positioning problem, see Figure 4.1. Any errors, orbital and atmospheric, experienced by the reference and rover receivers are correlated, therefore they will be significantly reduced after single differential processing. The reduction of these errors is dependent upon the baseline distance between the rover and reference station. Regarding the satellite clock error, it is known to be highly stable (Kaplan, 1996), thus, its value tends to cancel out for observations that are differenced at

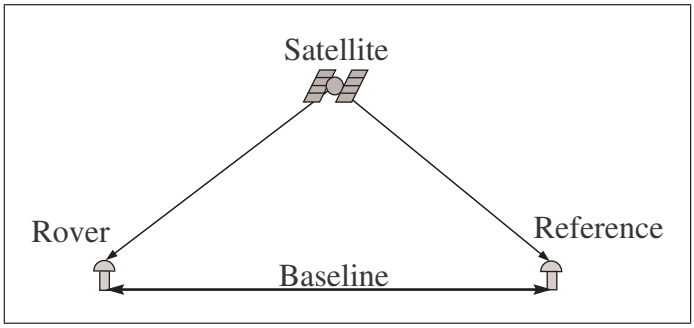


Figure 4.1: *Single difference concept.*

approximately the same time and is therefore removed in Equation (4.8). The only draw back of SD measurements is that the amplitude of uncorrelated errors, such as multipath and receiver noise, are enhanced by a factor of $\sqrt{2}$.

The parameter that requires special attention is the receiver clock error. The fact that the SD process does not removes this term means that it must be estimated. This becomes a problem for carrier phase observation, as it also contains an ambiguity term, which is also an unknown. In the least squares adjustment, the receiver clock error and integer ambiguity are actually inseparable; meaning that the estimation process cannot isolate the two terms, and is therefore unable to resolve both terms at the same time. Due to this reason, SD is usually limited to positioning with pseudoranges, where the measurements are not ambiguous.

4.3.2 Double Difference

As discussed in the previous section, the relative receiver clock error, ΔdT, in Equation (4.8) still remains after the SD process. Knowing that the relative receiver clock errors of the SD measurements are actually common for all satellites at the same epoch, it can be eliminated by forming between satellite differences, see Figure 4.2.

To form these Double Difference (DD) observations, SD carrier phase measurements

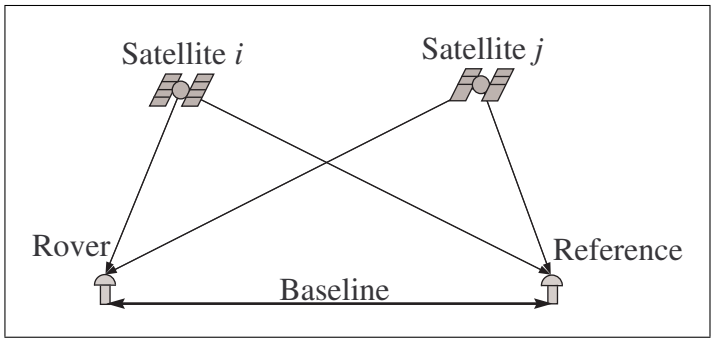


Figure 4.2: *Double difference concept.*

from satellite i and j are differenced at the same epoch.

$$\begin{aligned} \Delta\nabla\phi &= \Delta\phi^i - \Delta\phi^j \\ &= [\Delta\nabla\rho + \Delta\nabla d\rho - \Delta\nabla d_{\text{ion}} + \Delta\nabla d_{\text{trop}}]\,\lambda^{-1} + \Delta\nabla N + \Delta\nabla\varepsilon_\phi, \end{aligned} \quad (4.9)$$

where $\Delta\nabla$ is the DD operator. DD measurements possess many advantages over undifferenced measurements. The first advantage is that the DD process can remove all receiver and satellite clock errors. Second, orbital and atmospheric errors can be significantly reduced for short baselines. With most error sources reduced or eliminated, the resolution of the integer ambiguities can be carried out more effectively. However, DD also has its disadvantage. The DD noise $\Delta\nabla\varepsilon_\phi$ is twice the level compared to undifferenced measurements (Misra & Enge, 2001), hence making multipath and receiver noise more significant in DD measurements.

4.4 Linear Phase Combinations

4.4.1 Widelane Observable

The resolution of integer ambiguities is dependent upon the wavelength of the carrier. The longer the wavelength, the easier it is to resolve ambiguities. With dual-frequency measurements, an observable with a much longer wavelength known as the Wide-Lane (WL) can be derived. In order to form WL measurements, the L_2 carrier phase measurements is combined with the L_1 carrier to form a linear combination.

$$\phi_{\mathrm{WL}} = \phi_{\mathrm{L_1}} - \phi_{\mathrm{L_2}}$$

The wavelength of the WL observable is 86 cm which is significantly longer than the wavelength of L_1 (Liu, 2003). With a longer wavelength, WL measurements are much more resistant to position errors compared to L_1. For example, it requires only 19 cm of position error to introduce one cycle of error for L_1, and it requires 86 cm for WL. As for the ionospheric error of WL in units of cycles, it can be expressed as

$$\begin{aligned} \frac{I_{\mathrm{WL}}}{\lambda_{\mathrm{WL}}} &= \left(\frac{\lambda_{\mathrm{L_2}}}{\lambda_{\mathrm{L_1}}} - 1\right) \frac{d_{\mathrm{ion}}}{\lambda_{\mathrm{L_1}}} \\ &= \frac{17}{60} \frac{d_{\mathrm{ion}}}{\lambda_{\mathrm{L_1}}}. \end{aligned} \tag{4.10}$$

The ionospheric error of the WL is only a fraction of its wavelength, which means that WL measurements are much more resistant to ionospheric effects. Therefore, the resolution of WL ambiguities is much more reliable than L_1. Although WL ambiguities can be resolved more reliably, the ionospheric error in metres is actually amplified. By multiplying the WL ionospheric error in units of cycles with its wavelength, the ionospheric error in metres is $\frac{77}{60} d_{\mathrm{ion}}$, which is greater than the error on L_1.

Furthermore, the noise level of WL measurements is amplified compared to the raw L_1 and L_2 carriers. Thus, the position estimates derived from WL only measurements have larger errors compared to the L_1 measurements.

Other linear combinations, such as Narrow-Lane (NL) and Ionosphere-Free (IF) measurements, are not used and will not be described in detail. Table 4.2 summarizes the magnitude of the ionospheric errors and noise values for the different phase combinations.

Table 4.2: *Properties of different phase combinations (Liu, 2003).*

Observation	Wavelength [cm]	Ambiguity	Ionospheric Error [m]	Ionospheric Error [cycles]	Noise [cm]	Noise [cycles]
L_1	19.03	N_{L_1}	d_{ion}	$\frac{d_{\text{ion}}}{\lambda_{L_1}}$	19σ	σ
L_2	24.42	N_{L_2}	$\left(\frac{77}{60}\right)^2 d_{\text{ion}}$	$\frac{77}{60}\frac{d_{\text{ion}}}{\lambda_{L_1}}$	24σ	σ
WL	86.19	$N_{L_1} - N_{L_2}$	$-\frac{77}{60}d_{\text{ion}}$	$\frac{17}{60}\frac{d_{\text{ion}}}{\lambda_{L_1}}$	122σ	$\sqrt{2}\sigma$
NL	10.70	$N_{L_1} + N_{L_2}$	$\frac{77}{60}d_{\text{ion}}$	$\frac{137}{60}\frac{d_{\text{ion}}}{\lambda_{L_1}}$	15σ	$\sqrt{2}\sigma$
IF	48.44	$N_{L_1} - \frac{\lambda_{L_1}}{\lambda_{L_2}}N_{L_2}$	0	0	60σ	1.26σ

Chapter 5

Software Simulator

SIMGNSSII^{TM} is a software simulator developed by the Position, Location And Navigation (PLAN) Group in order to augment the use of expensive and complicated hardware simulators. This simulator is programmed in C++ and is capable of simulating code and carrier phase measurements for users in both static and kinematic modes (Cannon & Lachapelle, 2004). The algorithms used in this software simulator are discussed below.

5.1 Observation Model

The simulation of GPS measurements is done on an epoch-by-epoch, and user-by-user, basis. For each epoch and for each user, the positions and velocities of the satellites in sight are computed based on a set of almanacs. The almanac files are provided by the U.S. Coast Guard Navigation Center (USCG, 2002). Since almanac files contain only the main orbital parameters, e, t_{oe}, i_0, $\sqrt{a}$, $\dot{\Omega}$, Ω_0, ω and M_0, the remaining ephemeris parameters are set to zero,which means no perturbations are considered. Furthermore, no orbital errors are included in the simulated orbits, as orbital errors are directly added into the code and the carrier phase measurements if they are to be considered.

The user defines the receivers positions and velocities in the trajectory files as inputs, which allows the computation of the true ranges and rates between satellites

and users. The code and carrier phase measurements are generated by contaminating the true range with different error sources based on observation equations, which are repeated here for convenience.

$$p = \rho + d\rho + d_{\text{ion}} + d_{\text{trop}} + c(dt - dT) + \varepsilon_p \tag{5.1a}$$

$$\phi = [\rho + d\rho - d_{\text{ion}} + d_{\text{trop}} + c(dt - dT)]\,\lambda^{-1} + N + \varepsilon_\phi. \tag{5.1b}$$

All the errors are simulated based on mathematical models and will be discussed in the following sub-sections. The satellite and receiver clock errors are, however, not simulated and are assumed to be zero (Dong, 2004). This assumption will not affect the results, since differential techniques can completely remove the clock errors anyway. The final outputs of SIMGNSSII$^{\text{TM}}$ are a ephemeris file and an observation file that contain all the simulated measurements. The simulation procedures of SIMGNSSII$^{\text{TM}}$ are shown in Figure 5.1.

5.2 Orbital Error Model

As mentioned in the previous section, the GPS orbits are simulated based on almanacs and without any perturbations. Therefore, simulated orbital errors are directly included into the measurements to replicate the mis-modelling of GPS orbits. The simulation is based on the statistical properties of actual orbital errors extracted by comparing satellite positions derived from the broadcast ephemeris with the precise orbit. These orbital information can be downloaded from the International GNSS Service. The accuracy of precise orbits is less than 5 cm (IGS, 2005), which is extremely accurate compared to the broadcast ephemeris. According to the analysis

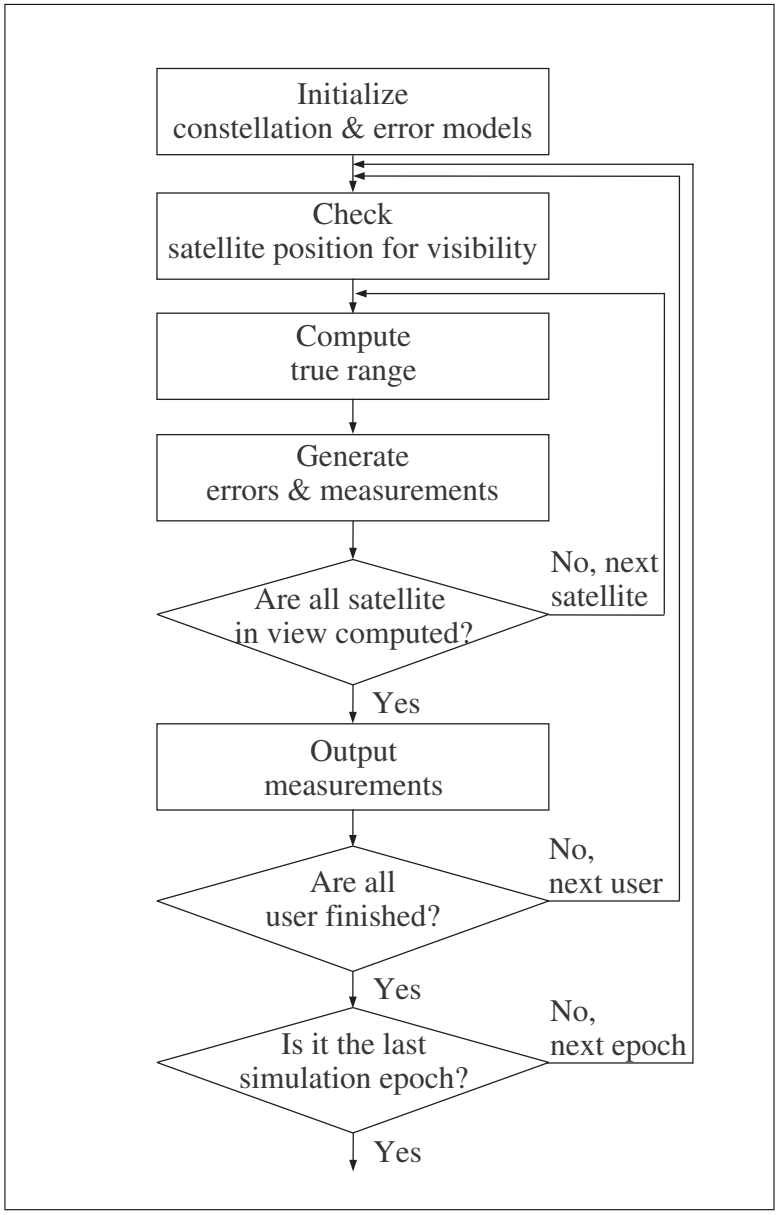


Figure 5.1: SIMGNSSII^{TM} *software simulator (Cannon & Lachapelle, 2004).*

of Luo (2001), orbital errors change slowly over time and consist of long term correlations. Furthermore, the probability distribution of the error is extremely close to Gaussian. Characterization of the temporal variations of orbital errors requires a correlation function. However, due to computation difficulties, the determination of the correlation function is rarely done in the time domain for random sequences that are too long. Instead, the Welch's averaged periodogram method (Welch, 1967) is first used to compute the power spectral density of the error and then the correlation function is determined using an inverse Fourier transform (Luo, 2001). The fact that orbital errors are approximately Gaussian distributed enable their generation with a auto-regressive model. The coefficients of the model are determined using the empirical correlation function. Furthermore, to increase the output rate of the model, a ninth-order Lagrange interpolator is added to the generator.

Examples of simulated orbital errors that are added to the GPS measurements are plotted in Figure 5.2. The simulation is conducted for a spacecraft flying at LEO altitude. The RMS error is approximately 2.4 m, which complies with the level specified by the IGS (2005). Though the simulation extends for only 6 hours, but large variation in the orbital errors are visible. Such a large error gradient is related to the extreme velocity of the spacecraft.

To analyze the orbital error between the spacecraft in an elliptical formation, the relative range error is illustrated in Figure 5.3. The dashed black line is the horizontal displacement of the elliptical formation. Note that the relative orbital error is correlated with the horizontal displacement, which demonstrates the spatial decorrelation property of the orbital error and its relation with the elliptical configuration.

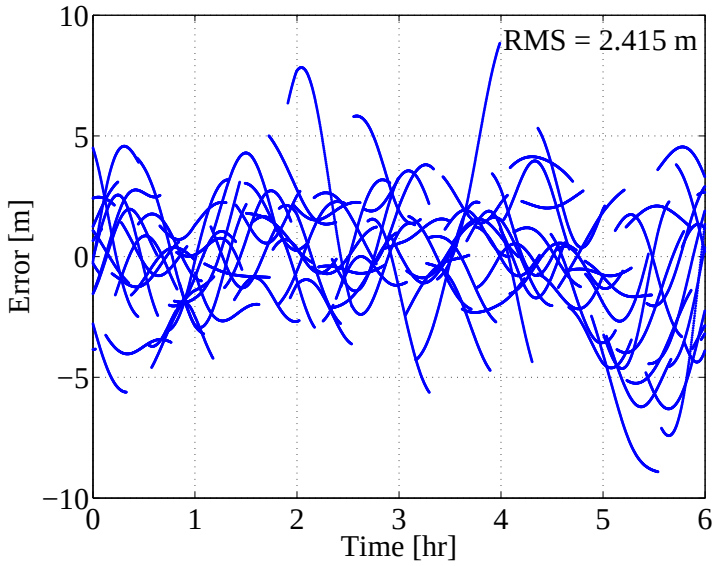


Figure 5.2: *Simulated orbital error for* LEO *spacecraft.*

5.3 Ionospheric Error Model

5.3.1 Vertical Total Electron Content Model

As mentioned in Section 4.2.2, the ionospheric errors are directly proportional to the TEC. In order to simulate ionospheric errors for the GPS measurements, the TEC must first be modelled. The model chosen uses a Spherical Harmonic (SH) expansion to represent the Vertical Total Electron Contect (VTEC) globally in the solar-geomagnetic reference system (Luo, 2001).

$$\text{VTEC}(\beta, s) = \text{SF} \sum_{n=0}^{n_{\max}} \sum_{m=0}^{n} \bar{P}_{nm}(\sin\beta)(a_{nm}\cos ms + b_{nm}\sin ms) \tag{5.2}$$

where β is the geocentric latitude of the pierce point, s is the solar-geomagnetic longitude of the pierce point, SF is the scale factor that controls the intensity of the ionosphere, $\bar{P}_{nm}(\sin\beta)$ is the normalized Legendre function of degree m and order n,

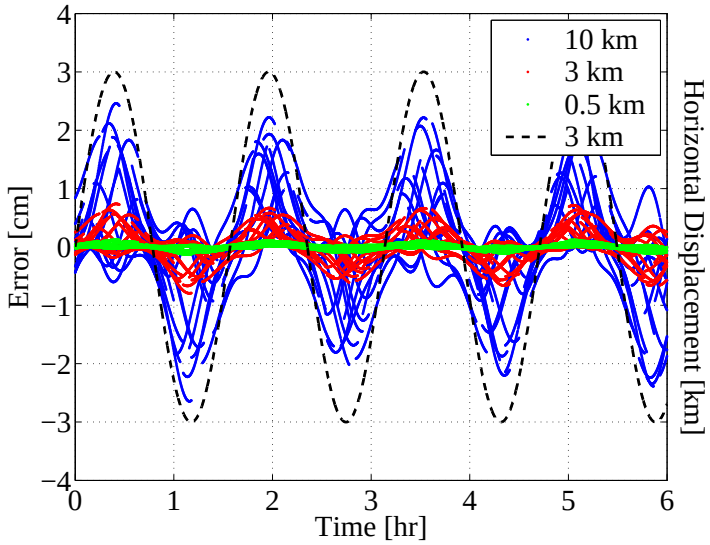


Figure 5.3: *Relative orbital errors for 2-by-1 elliptical formations.*

and a_{nm}, b_{nm} are the SH coefficients. The SH coefficients of the VTEC are available in Global Ionosphere Map (GIM) files generated by the Centre for Orbit Determination in Europe CODE (2006). They are produced on a daily basis using data compiled from approximately 200 GPS and GLONASS sites. The GIM can provide SH expansion up to degree and order 15. To separate the seasonal effect of the ionosphere, four coefficient files are generated by averaging all the GIM files to the corresponding seasons.

For terrestrial applications, where receivers are on the Earth's surface, the SF can be taken to be 1 since the VTEC of the GIM are derived from ground stations. However, the same SF is not adequate for formation flying spacecraft at LEO altitudes, since the GPS signal received in space experiences much less impact from the ionosphere. To model the ionospheric errors in a spaceborne environment, the vertical structure of the ionosphere is required. Such knowledge is important for the

determination of the correct SF for upward continuation of the VTEC from zero to LEO altitude. Richert (2003) has presented a normalized VTEC model as a function of the height of projectiles. For formation flying spacecraft 500 km above the Earth's surface, the normalized VTEC value is approximately 0.5, which is half of the TEC at ground level. This value is reasonable since a spacecraft is above the F_2 layer of the ionosphere, which is where the free electrons are most concentrated.

The spatial resolution of the VTEC model is limited by the maximum degree and order of the coefficients available.

$$\Delta\beta = \frac{2\pi}{n_{\max}} \tag{5.3a}$$

$$\Delta s = \frac{2\pi}{m_{\max}} \tag{5.3b}$$

where $\Delta\beta$ and Δs are the resolutions in the latitude and longitude directions. For the GIM, the maximum degree and order are 15, which corresponds to a spatial resolution of 24°. Such a spatial resolution is certainly not enough to obtain satisfactory regional ionospheric error because of the absence of high degree and order SH coefficients. To overcome this problem, high frequency variations are included by adding higher degree and order components into the SH model, see Table 5.1.

Table 5.1: *High degree and order* SH *coefficients of the* VTEC *model.*

degree (n)	order (m)	coefficient
50	25	1
100	50	1
100	-50	1
200	100	1
200	-100	1

Although the spatial resolution is enhanced, the temporal resolution still remains

unchanged. The fixed SH coefficients represent a static ionosphere in the solar-geomagnetic reference system. The dynamics are very low because the variation is only from β and s at different simulation epochs. To simulate short term ionospheric error variations, it is necessary for the temporal resolution to improve. This can be achieved by updating the high degree and order SH coefficients. Using a 1st order Gauss-Markov model, the temporal resolution is improved by modifying the high degree and order SH coefficients (Dong, 2004).

$$n_k = (1-\theta)n_{k-1} + \theta w \tag{5.4a}$$

$$X_k = e^{\frac{-\Delta t}{\tau}} X_{k-1} + n_k \tag{5.4b}$$

where k is the index for the current epoch, w is Gaussian white noise with a defined variance, θ is the noise smoothing factor, Δt is the sampling time interval, τ is the correlation time and X is the magnitude of the updated SH coefficients. With temporal resolutions enhanced, the VTEC model is now much more practical.

5.3.2 Grid Network

With the SH coefficients available, the VTEC at any pierce point can be determined by SH synthesis (5.2). However, this method can cause severe computational difficulties, since for every point, the normalized Legendre function needs to be recomputed for all degrees and orders. To overcome this problem a TEC network is first created and then the VTEC at pierce points can be estimated by planar interpolation. This method can limit the number of points computed by synthesis, hence, reducing the computational load. With the TEC network available, the VTEC at pierce points are interpolated using the four closest grid points. The interpolation of the VTEC can

be computed using the following equation:

$$\text{VTEC} = \sum_{i=1}^{4} \text{W}_i \cdot \text{VTEC}_i, \tag{5.5}$$

where W_i are the spatial weighting factors for each network point. In this algorithm, weights are determined based on the bilinear interpolation weighting function (RTCA, 1998).

5.3.3 Ionospheric Model

To compute the ionospheric error, the Slant Total Electron Content (STEC) is required. The ionospheric model is therefore introduced to convert VTEC to STEC. The model assumes that the ionosphere is a single layer with all free electrons distributed in a spherical shell that is of infinitesimal thickness, see Figure 5.4. The

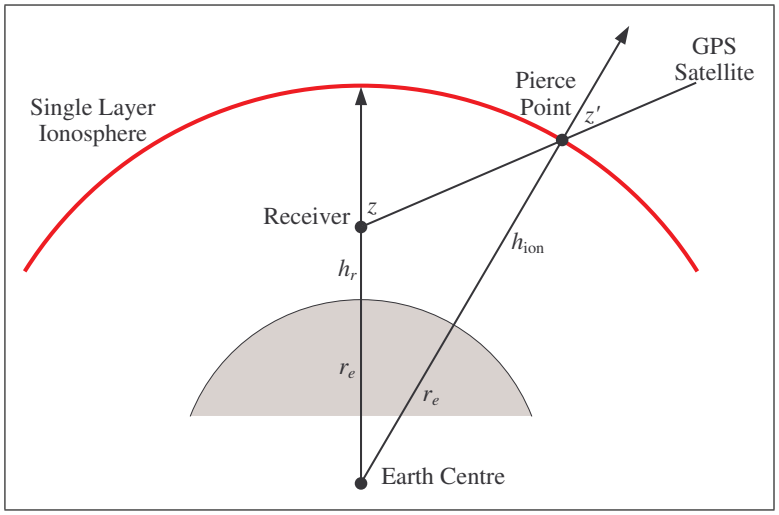


Figure 5.4: *Single layer ionosphere model (Moon, 2004).*

GPS signal path length is dependent upon the elevation of the satellites. The lower the elevation, the longer the path length and the higher the STEC. Assuming that

there are no lateral electron gradients, the STEC can be computed by multiplying the VTEC with a mapping function.

$$\text{STEC} = \frac{1}{\cos z'}\text{VTEC} \tag{5.6}$$

where z' is the zenith angle at the pierce of single layer ionosphere, which can be expressed in terms of the zenith angle at the height of the receiver z using the law of sines.

$$\frac{\sin z'}{r_e + h_r} = \frac{\sin z}{r_e + h_{\text{ion}}}, \tag{5.7}$$

where h_r is the height of the receiver, r_e is the mean radius of the Earth and h_{ion} is the height of the single layer ionosphere.

For terrestrial applications, the height of the ionosphere is usually set to the F_2 layer, 450 km above the Earth's surface, where the density of free electrons is highest. However, this cannot be applied to formation flying spacecraft orbiting at 500 km altitude. To apply the existing model to spacecraft at LEO altitudes, changing the height of the single layer ionosphere is necessary. Assuming that the TEC above 1000 km is negligible, the height of the ionosphere can be determine as a function of the spacecraft's altitude.

$$h_{\text{ion}} = \frac{h_r + 1000}{2} \tag{5.8}$$

which is the average height between the spacecrafts and the maximum ionospheric height. This assumes that there are no vertical electron gradients within the ionosphere; although this is not true, it does provide a reasonable approximation for determining the ionospheric errors at a spaceborne environment, since it is used only for computing the STEC.

An example of the ionospheric error simulated for a spacecraft flying in equatorial orbit and at 485 km altitude is plotted in Figure 5.5. The initial position of the spacecraft is at Greenwich meridian and the initial time is midnight local time. The

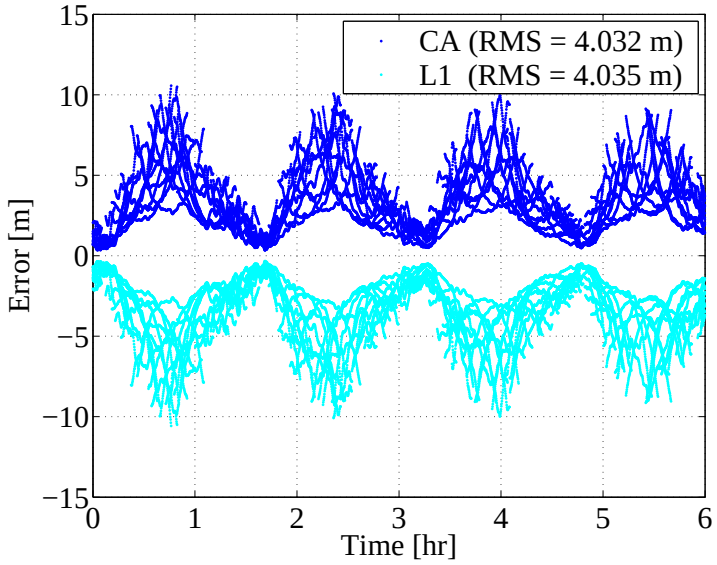


Figure 5.5: *Simulated code and carrier ionospheric errors for* LEO *spacecraft.*

figure shows the ionospheric effect on both C/A code and L_1 carrier phase measurements. Though the simulation time is only six hours, but there are four peaks visible from the plot. The reason is due to the extremely velocity of LEO spacecraft, since the orbital period is approximately 1.5 hours.

To analysis the relative ionospheric effect, the ionospheric phase advances of L_1 are plotted in Figure 5.6 for elliptical formations with 0.5, 3 and 10 km baselines. The errors are simulated with a low ionospheric condition where the average relative error level is approximately 2 ppm. The dashed black line plotted is the vertical displacement of a 10 km 2-by-1 elliptical formation and therefore, the vertical displacement only reaches 5 km.

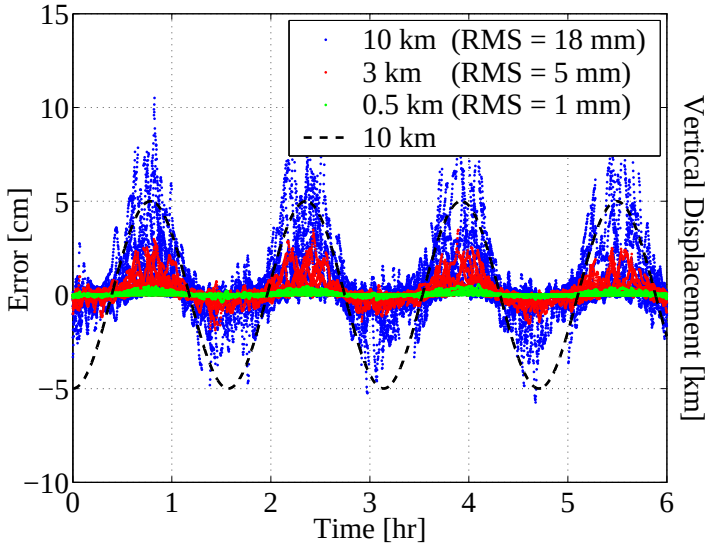


Figure 5.6: *Relative* L_1 *ionospheric errors for 2-by-1 elliptical formations.*

The plot indicates that the the relative ionospheric error is correlated with vertical motion. This is because the absolute and relative motion of spacecraft have a common orbital frequency, therefore the formations always enters the maximum ionospheric zone when vertical displacements largest. Examining the plot closely shows that the ionospheric peaks and dips are actually shifted slightly, which is related to the ionization delay. For terrestrial applications, the ionospheric peak is approximately two hours delayed from the local noon. However, due to the extreme velocity of LEO spacecraft, the delay is compressed to only a fraction of a hour. This explains the minor shift of the ionospheric peak and dip shown in the plot.

Simplifications and assumptions have been made to reduce the simulation procedure's complexity. Only a 2-dimensional ionosphere model has been used and the vertical gradient has been neglected. Therefore, the simulated errors shown are some what incorrect due to the simplification. For the scenario shown above, the

relative errors expected during the peaks should be slightly higher due to the vertical displacement. Nevertheless the simulated errors are still within a reasonable range.

5.4 Multipath Error Model

5.4.1 Multipath Environment Model

In a spaceborne environment, the only multipath expected is the locally reflected signals since there are no other potential reflectors that are close enough to induce significant multipath. Only short local multipath from the spacecraft themselves is expected to interfere with the direct signal. Since multipath is highly dependent upon the environment, a multipath environment model is developed to simulate local reflected signals. The model used is a simplified version of the sophisticated model developed by Ray (2000). In the simplified model, the only reflector is the infinite ground plane, which is similar to the highly reflective solar panels on spacecraft, see Figure 5.7. In this configuration, the antenna is capable of receiving one and only

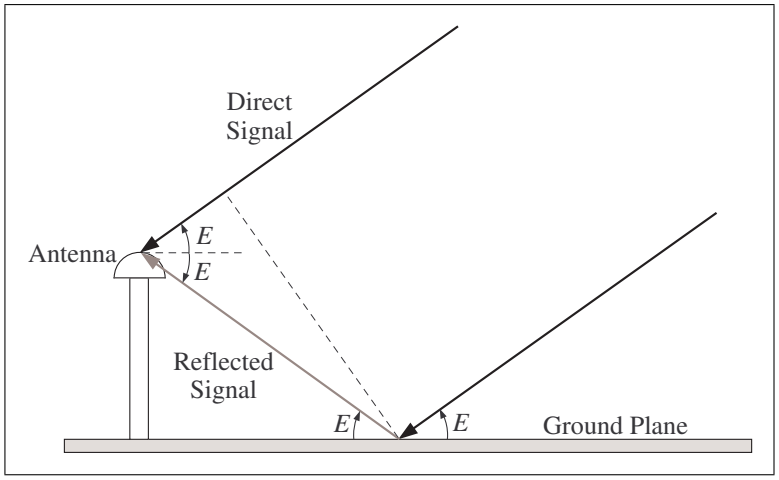


Figure 5.7: *Multipath environment model (Dong, 2004).*

one reflected signal from the ground, which can be expressed as

$$s(t) = \sum_{i=0}^{1} \alpha_i c(t - \tau_i) \cos(\omega t + \gamma_i), \tag{5.9}$$

where $i = 0$ is the direct signal, $i = 1$ is the reflected signal, α is the attenuation coefficient with $\alpha_0 = 1$, c is the code being used to produce pseudorange, τ is the code phase delay, ω is the frequency of the carrier signal and γ is the carrier phase shift. In this model, the attenuation coefficient is a normally distributed random variable with a mean value of 0.4 and a standard deviation of 0.2. To compute the path delay, signals arriving at the reflecting point are assumed to be parallel with the direct signal, since the satellites are far away. Thus, the path delay and phase shift of the reflected signals are only dependent upon the elevation angle of the satellite E and the height of the antenna h.

$$\tau_1 = 2h \sin E \tag{5.10a}$$

$$\gamma_1 = \frac{2\pi\tau}{\lambda} \tag{5.10b}$$

Due to the simplicity of this model, the elevation angles are nearly the same for spacecraft in configurations with reasonable inter-satellite distances ($< 100\,\text{km}$). A common elevation angle would produce very similar multipath environments between the spacecraft. To create differences in their environment, different antenna heights are assigned to each of the spacecraft. This ensures that the multipath between spacecrafts are totally uncorrelated.

5.4.2 Receiver Tracking Loops

With the multipath environment defined, the multipath error can now be determined by analyzing the tracking loops of the receiver. There are two distinct tracking loops

for each channel within the receiver. Code tracking is implemented as a feedback control loop, called a Delay Lock Loop (DLL), which aligns the code replica with the code of the incoming signal. As for phase tracking, the Phase Lock Loop (PLL) is used to match the phase and the frequency of the incoming signal. The two lock loops are coupled with one another to generate measurements and to extract the navigation message from the incoming signal.

Within the DLL, early, prompt and late versions of the code are generated (Julien, 2005). The duration between the early and late codes are dependent upon the correlator spacing. Here, a narrow correlator with 0.05 chips spacing is used for conducting the integration. In the PLL, a reference carrier signal is generated for Doppler removal from the in-phase and quadrature-phase of the incoming signal. These two signals are then integrated with the early, prompt and late versions of the code to generate six correlation values (Ray, 2000).

$$\text{IE} = \frac{1}{2}\sum_{i=0}^{1}\alpha_i R(\hat{\tau}-\tau_i+T_d)\cos(\gamma_i-\hat{\gamma}) \tag{5.11a}$$

$$\text{IP} = \frac{1}{2}\sum_{i=0}^{1}\alpha_i R(\hat{\tau}-\tau_i)\cos(\gamma_i-\hat{\gamma}) \tag{5.11b}$$

$$\text{IL} = \frac{1}{2}\sum_{i=0}^{1}\alpha_i R(\hat{\tau}-\tau_i-T_d)\cos(\gamma_i-\hat{\gamma}) \tag{5.11c}$$

$$\text{QE} = \frac{1}{2}\sum_{i=0}^{1}\alpha_i R(\hat{\tau}-\tau_i+T_d)\sin(\gamma_i-\hat{\gamma}) \tag{5.11d}$$

$$\text{QP} = \frac{1}{2}\sum_{i=0}^{1}\alpha_i R(\hat{\tau}-\tau_i)\sin(\gamma_i-\hat{\gamma}) \tag{5.11e}$$

$$\text{QL} = \frac{1}{2}\sum_{i=0}^{1}\alpha_i R(\hat{\tau}-\tau_i-T_d)\sin(\gamma_i-\hat{\gamma}) \tag{5.11f}$$

where IE, IP, IL, QE, QP and QL are the In-phase Early, In-phase Prompt, In-Phase Late, Quadrature-phase Early, Quadrature-phase Prompt and Quadrature-

phase Late correlation values respectively, R is the correlation function of the code, $\hat{\tau}$ is the tracking loop estimate of the code delay of the incoming signal, $\hat{\gamma}$ is the tracking loop estimate of the phase shift and T_d is the chip spacing of the correlator.

With the correlation values available, code and phase tracking errors can now be determined by the discriminator. A dot-product non-coherent discriminator is adopted by the DLL and an arc-tangent discriminator is used by the PLL (Ward, 1996).

$$D_{\text{DLL}} = \text{IP}(\text{IE-IL}) + \text{QP}(\text{QE-QL}) \tag{5.12a}$$

$$D_{\text{PLL}} = \arctan\left(\frac{\text{QP}}{\text{IP}}\right) \tag{5.12b}$$

where D is the discriminator function. In the tracking loop, both the DLL and PLL try to align with the incoming signal by minimizing the discriminator value to zero. If the environment is ideal where no multipaths exist, then the estimate of the code delays and phase shifts will be perfect. However, due to multipath, errors are introduced into the measurements. The phase multipath error can be determined analytically by letting the discriminator value in Equation (5.12b) be equal to zero.

$$\hat{\gamma} = \arctan\left(\frac{\alpha_1 R(\hat{\tau} - \tau_1)\sin\gamma_1}{R(\hat{\tau}) + \alpha_1 R(\hat{\tau} - \tau_1)\cos\gamma_1}\right) \tag{5.13}$$

To simplify the computation, the estimate of the code delays are assumed to be delays of the direct signal which are set to zero. As for the code multipath errors, an iteration method is used to find the specific code error $\hat{\tau}$ that makes the discriminator value in Equation (5.12a) equal to zero. Simulated code and phase multipath errors are shown in Figures 5.8 and 5.9, which are similar to the error levels investigated by Montenbruck & Kroes (2003) and Kroes (2006).

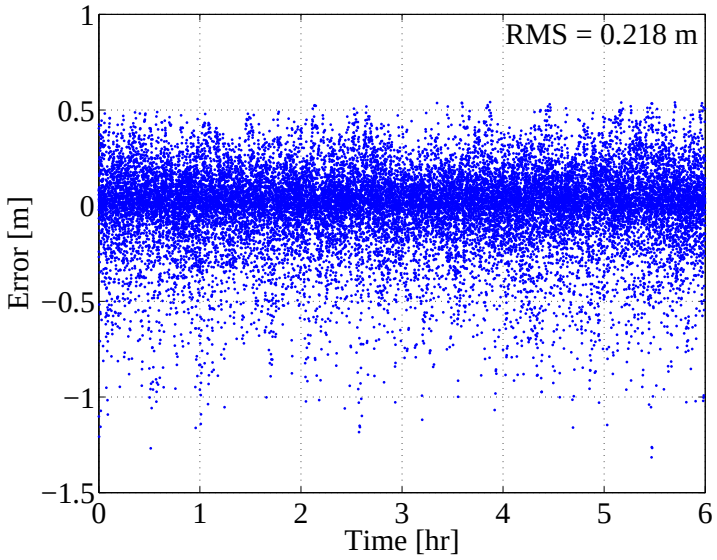


Figure 5.8: *Simulated* C/A *code multipath errors.*

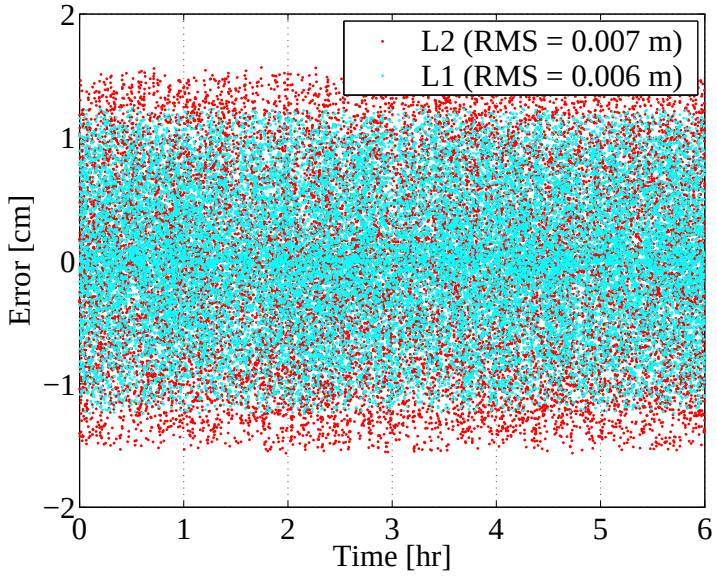


Figure 5.9: *Simulated* L_1 *and* L_2 *carrier multipath errors.*

5.5 Receiver Noise

Receiver noise is mainly caused by thermal noise, dynamic stress and oscillator stability in the tracking loop. It can be isolated from all other errors by using a zero baseline test, where two receivers are connected to the same antenna and differential techniques are used to estimate the variance of the receiver noise.

Receiver noise is simulated by using a normalized random Gaussian noise generator. The measurement noise is obtained by scaling the normalized noise with the corresponding standard deviation. The noise level of the C/A code-derived pseudorange is approximately 5 cm. As for the noise level of the phase measurements, Montenbruck (2003) conducted a zero baseline test with a NovAtel OEM4-G2L receiver and estimated the RMS error to be 0.8 mm for L_1 carrier, 1 mm for L_2 carrier, and 1.5 cm/s for the L_1 Doppler measurement. These noise levels are within reasonable ranges and similar to the values of other receivers (e.g. NovATel GPSCard^{TM}), therefore they are adopted for this research. Examples of the simulated C/A code and L_1 carrier phase noise are shown in Figures 5.10 and 5.11.

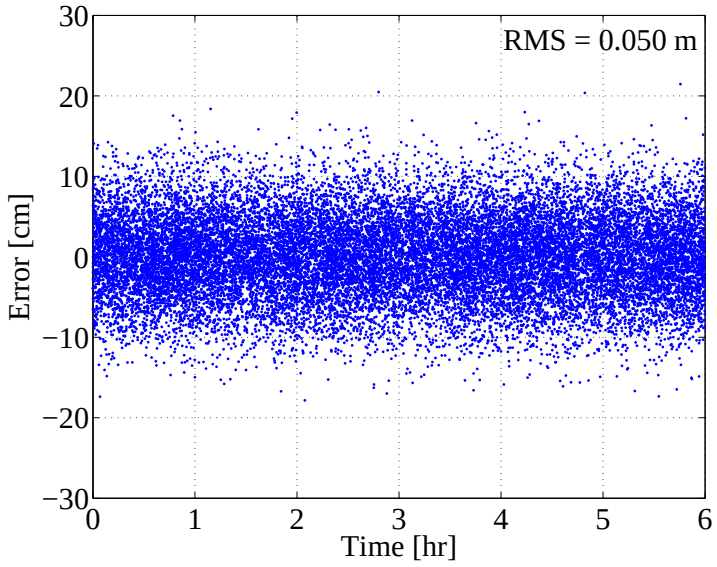


Figure 5.10: *Simulated* C/A *code noise.*

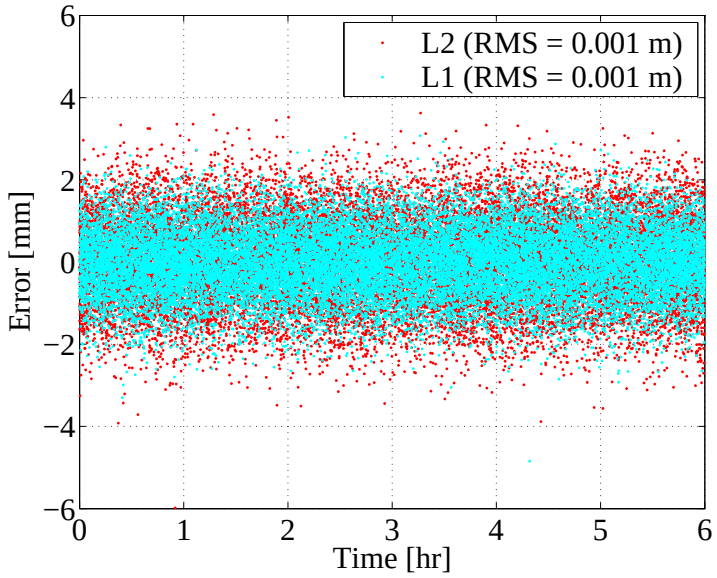


Figure 5.11: *Simulated* L_1 *and* L_2 *carrier noises.*

Chapter 6

Relative Navigation Algorithms

FLYKIN+^{TM} is a C++ program developed by the Positioning, Location and Navigation (PLAN) Group for the purpose of processing pseudorange, carrier phase and/or Doppler measurements for relative navigation. However, the original software limits the reference station to be a static receiver, which is not applicable to formation flying. To overcome this problem, the Moving Base Station (MBS) ability is modified and used to allow for relative positioning between two moving receivers (Crawford, 2005). Furthermore, a new dynamic model based on the Hill Equations of motion is implemented for the purpose of spacecraft formation flying. In this chapter, details regarding the algorithms will be discussed.

6.1 Measurement Model

The measurement model used for relative navigation is based on Double Difference (DD) observations discussed in Section 4.3.2. The feasibility of both single and double difference carrier phase measurements has also been demonstrated by Busse (2003). The advantage to using DD measurements is the reduced orbital and atmospheric errors, and the elimination of receiver and satellite clock errors. The downside of DD measurements is that uncorrelated errors, such as multipath and noise, are amplified by a factor of 2. By using the DD observation model, all residual errors are lumped together and assumed to be white noise. The DD measurement model includes C/A

code derived pseudorange, L_1 carrier phase, Widelane (WL) carrier phase and L_1 Doppler measurements.

$$\Delta\nabla p_{\mathrm{CA}} = \Delta\nabla\rho + \Delta\nabla\varepsilon_p \tag{6.1}$$

$$\Delta\nabla\phi_{\mathrm{L1}} = \frac{\Delta\nabla\rho}{\lambda_{\mathrm{L1}}} + \Delta\nabla N_{\mathrm{L1}} + \Delta\nabla\varepsilon_{\mathrm{L1}} \tag{6.2}$$

$$\Delta\nabla\phi_{\mathrm{WL}} = \frac{\Delta\nabla\rho}{\lambda_{\mathrm{WL}}} + \Delta\nabla N_{\mathrm{WL}} + \Delta\nabla\varepsilon_{\mathrm{WL}} \tag{6.3}$$

$$\Delta\nabla\dot{\phi}_{\mathrm{L1}} = \frac{\Delta\nabla\dot{\rho}}{\lambda_{\mathrm{L1}}} + \Delta\nabla\varepsilon_{\dot{\phi}} \tag{6.4}$$

The use of pseudorange measurements is only before the resolution of ambiguities, as its contribution to the navigation solution becomes insignificant when the full potential of L_1 carrier phase measurements are exploited. The use of WL measurements is to take advantage of its extended wavelength to improve the ambiguity resolution process. As for the L_1 Doppler measurements, they are included to facilitate the estimate of the relative velocity. Different combination of these measurements will be thoroughly analyzed in Section 7.1.5 to determine the optimal observation set to be used for relative navigation.

6.2 Kalman Filter

The Kalman Filter (KF) is commonly used to process differential GPS measurements for optimal estimation of the relative position and velocity of formation flying spacecraft. The main advantage of using the KF for navigation purposes is that it is a simple recursive algorithm that delivers optimal statistical estimation of the state parameters, which is particularly suitable for real-time applications. Another advantage is that even in the absence of GPS measurements, the filter is still able to

operate hence, providing robustness to the estimation. The recursive algorithm of the classical KF can be separated into two recursive steps which are the prediction stage and the correction stage (Axelrad & Enge, 1996).

6.2.1 Prediction Stage

Kinematic and stochastic models are the mathematical expressions used for predicting the state parameters within a dynamic system. Assuming that the system is linear and time-invariant, then both deterministic and stochastic state variables of the system can be described by linear ordinary differential equations. By using a state space model, any high order dynamic system can be represented by a first order differential equation.

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{w}(t) \quad \mathbf{Q}(t) \tag{6.5}$$

where $\mathbf{x}(t)$ is the continuous state vector as a function of time t, $\mathbf{F}$ is the matrix of the dynamic system, $\mathbf{G}$ is the coefficient matrix of the random forcing function, $\mathbf{w}(t)$ is the white noise vector and $\mathbf{Q}(t)$ is the spectral density matrix associated with the white noise. Equation (6.5) consists of a functional model of the system and a stochastic model for describing the uncertainties associated with the model. Solving the first order differential equation over time increments leads to a discrete form of the system,

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{w}_k \quad \mathbf{Q}_k \tag{6.6}$$

where $\mathbf{\Phi}$ is the transition matrix for propagating the state vector from time k to $k+1$, $\mathbf{w}_k$ is the process noise with its elements being white sequences and $\mathbf{Q}_k$ is the

covariance matrix associated with the process noise. The stochastic information of the state vector can be computed using error propagation.

$$\mathbf{P}_{k+1} = \mathbf{\Phi}\mathbf{P}_k\mathbf{\Phi}^\mathsf{T} + \mathbf{Q}_k \tag{6.7}$$

where $\mathbf{P}$ is the covariance matrix of the state vector. Equations (6.6) and (6.7) are the core equations for the prediction stage of the KF.

The transition matrix can be determined using a Laplace transformation which requires the transfer function in the Laplace domain. According to Gelb et al. (1974), the transition matrix can also be expressed as the matrix exponential of the dynamic matrix.

$$\mathbf{\Phi} = e^{\mathbf{F}\Delta t} \tag{6.8}$$

where Δt is the sampling interval. Equation (6.8) can also be approximated by a Taylor series expansion (Gao, 2004).

$$\mathbf{\Phi} = \mathbf{I} + \Delta t\mathbf{F} + \frac{1}{2!}\Delta t^2\mathbf{F}\mathbf{F} + \dots \tag{6.9}$$

where $\mathbf{I}$ is the identity matrix. For a situation with low dynamics and short sampling intervals, it is sufficient to approximate the transition with a second order expansion. The propagation process also has knowledge on the stochastic properties of the associated white sequences. The uncertainty of the model is described by the covariance matrix $\mathbf{Q}$, which can be written in integral form (Brown & Hwang, 1997).

$$\mathbf{Q}_k = \int_0^{\Delta t}\int_0^{\Delta t} \mathbf{\Phi}(\xi)\mathbf{G}(\xi)E\left[\mathbf{w}(\xi)\mathbf{w}(\eta)\right]\mathbf{G}^T(\eta)\mathbf{\Phi}^T(\eta)d\xi d\eta \tag{6.10}$$

If the process noise is assumed to grow linearly over time and the system is linear time-invariant then Equation (6.10) can be approximated by the area of a trapezoid. For high order filters, where velocity or acceleration states are estimated, this

approximation is always pessimistic, therefore certain elements within the process noise matrix should be adjusted accordingly.

Propagation of Ambiguities

The dynamics of the ambiguity states can be modelled as random constants since their values remain constant over time when there is no loss of lock. The transition matrix for a random constant process is simply an identity matrix and a zero matrix for the corresponding process noise (Liu, 2003).

$$\mathbf{\Phi_N} = \mathbf{I} \tag{6.11}$$

$$\mathbf{Q_N} = \mathbf{0} \tag{6.12}$$

As for the orbit propagation models, they will thoroughly discussed in Section 6.3.

6.2.2 Correction Stage

To obtain optimal estimates for the state parameters, measurements are used to improve the predicted states in the correction stage. The relationship between the measurements and state parameters is related linearly through the design matrix (Cannon, 1991).

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \quad \mathbf{R}_k \tag{6.13}$$

where $\mathbf{z}$ is the measurement vector, $\mathbf{H}$ is the design matrix, $\mathbf{v}$ is the vector of measurement error assumed to be white sequence and $\mathbf{R}$ is the covariance matrix associated with the measurement noise. The design matrix represents the linear relationship between the measurement and the state vector. It can be computed by taking the

partial derivatives of the observation models $\mathbf{h}$ with respect to the state vector.

$$\mathbf{H} = \frac{\partial \mathbf{h}(\mathbf{x_0})}{\partial \mathbf{x_0}} \tag{6.14}$$

where $\mathbf{x}_0$ is the initial best estimate of the state vector. The adjustment to the state vector is expressed as

$$\mathbf{x}_k^+ = \mathbf{x}_k^- + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k^-), \tag{6.15}$$

where $\mathbf{x}^+$ is the *a posteriori* state vector, $\mathbf{x}^-$ is the *a priori* state vector provided by the prediction model and $\mathbf{K}$ is known as the Kalman gain, which acts as the scaling factor for the adjustment. A large Kalman gain means that the measurements acquired are more reliable than the prediction model, therefore, the weights will be shifted towards the innovation sequence $\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k^-$ and vice versa. The particular Kalman gain chosen must provide an updated estimate that is optimal. This optimality is defined as minimization of the mean-square of the estimation error (least-squares). Without showing the derivations, the expression for the optimum Kalman gain is

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^\mathsf{T}(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \mathbf{R}_k)^{-1}, \tag{6.16}$$

where $\mathbf{P}^-$ is the covariance matrix of the *a priori* state estimate, which is improved in the correction stage by the incoming measurements.

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-, \tag{6.17}$$

where $\mathbf{P}^+$ is the covariance matrix of the *a posteriori* state estimate.

Note that the measurement model of Equation (6.14) assumes a linear relationship between the observations and the state vector, which is generally not true for

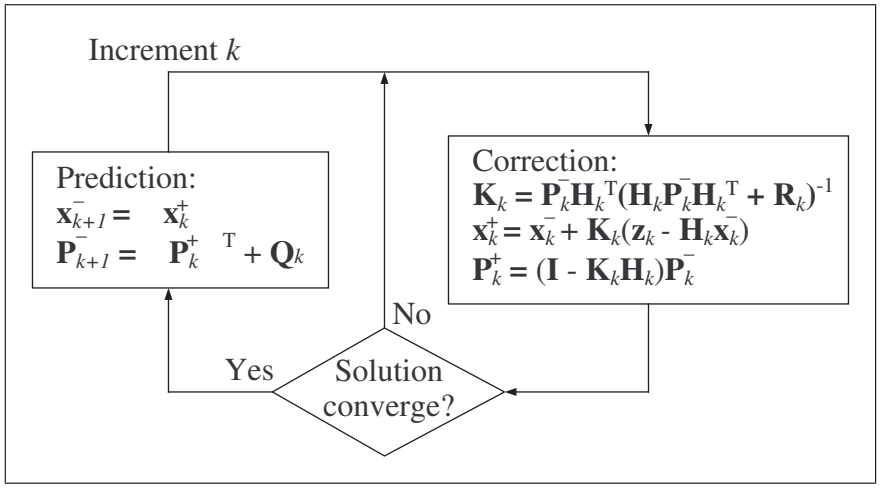


Figure 6.1: *Extended Kalman filter.*

geomatics applications. To overcome this problem, an Extended Kalman Filter (EKF) is implemented to linearize the measurement model about the state estimate iteratively (Montenbruck & Gill, 2000). Here, the correction stage is iterated until the state estimate converges to account for the non-linearity within the measurement model, see Figure 6.1.

The state vector of the EKF contains the following unknown parameters.

$$\mathbf{x} = \begin{bmatrix} \mathbf{r}_g \\ \dot{\mathbf{r}}_l \\ \mathbf{N} \end{bmatrix}, \tag{6.18}$$

where $\mathbf{r}_g$ is the absolute position vector of the rover receiver in a global geodetic frame which consists of the latitude, longitude and height (φ, λ, h), $\dot{\mathbf{r}}_l$ is the corresponding absolute velocity vector in a local geodetic frame consisting of the north, east and up velocities $(v_{\mathrm{N}}, v_{\mathrm{E}}, v_{\mathrm{U}})$ and $\mathbf{N}$ is the DD ambiguity vector which consist of ambiguities for each of the carrier phase measurements.

With measurements and state parameters defined, the design matrix can be com-

puted as the Jacobian of the observation equations.

$$\mathbf{H} = \left[\begin{array}{ccc} \dfrac{\partial \mathbf{h}}{\partial \mathbf{r}_g} & \dfrac{\partial \mathbf{h}}{\partial \dot{\mathbf{r}}_l} & \dfrac{\partial \mathbf{h}}{\partial \mathbf{N}} \end{array} \right], \tag{6.19}$$

which contains the partial derivatives of the measurement with respect to the filter states.

6.2.3 Moving Base Station

The first version of FLYKIN+$^{\text{TM}}$ was designed for land, shipborne and airborne applications in which the reference station is static. The original purpose of the software was implemented for precise positioning relative to the static reference station. However, the situation in space is totally different and requires some modifications to the algorithm in order to be applicable to formation flying.

For formation flying, the spacecraft are always in motion in order maintain to equilibrium with the gravitational potential. Therefore, there is no static reference station in this situation. To overcome this problem, the Moving Base Station (MBS) capability was added to the software, which allows relative navigation with non-static reference stations (Crawford, 2005).

In formation flying, the spacecraft acting as the reference station is commonly referred as the *chief* and other spacecraft within the configuration are known as *deputy*. Figure 6.2 is a representation of four deputy spacecraft revolving around a chief spacecraft in a 2-by-1 elliptical configuration. Within the EKF, the chief's position and velocity states are updated every epoch during the prediction stage. This information is delivered by another processing software called C^{3}NAVG$^{2\text{TM}}$. The program determines the absolute position and velocity independently from the

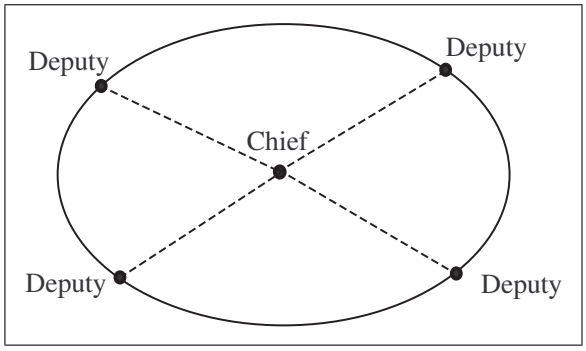


Figure 6.2: *Chief and deputy spacecrafts in 2-by-1 elliptical configuration.*

chief receiver's pseudorange and Doppler measurements by a least squares adjustment (Petovello et al., 2000). With this new implementation, FLYKIN+^{TM} is now applicable to satellite formation flying.

6.3 Orbit Propagation

Many kinds of propagation models have been implemented for the purpose of formation flying. Busse (2003) used a linearized relative dynamic model that considers only a central gravitational field. This model is similar to the Hill Equation of motion, but is referenced to the geodetic coordinate system. The advantage of such a linear model is a simple transition matrix that can be analytically derived for orbit propagation. A more sophistical propagation method has also been demonstrated by Kroes (2006). In his approach, instead of linearizing the relative dynamic model, the orbits of the spacecraft are numerically integrated through time to avoid any approximation. Though this method does provide more accurate *a priori* information for the filter, the complexity and processing load increase dramatically as well.

In this section, a rather simple kinematic model is introduced for orbit propa-

gation. To improve the orbit propagation performance, a relative dynamic model based on the Hill Equation of motion is also presented.

6.3.1 Kinematic Approach

Before deriving the transition matrix and process noise for the kinematic approach, an extra step is essential. This step is required in order to reduce the absolute velocity to a relative velocity for time propagation since the purpose of the procedure is to filter the baseline. The relative velocity cannot be computed by simply removing the chief velocity from the deputy since their velocities are represented in different local geodetic coordinate systems. To compute the relative velocity, scale factors must be applied to the deputy velocity to transform it into a common local frame with the chief. This transformation affects only the north and east velocities and the scale factor is the ratio of the their radii of curvature. Depending on the direction, the meridian and prime vertical radii of curvature in addition to the height should be used accordingly. To ensure proper implementation, the inversion of this transformation must also be applied after orbit propagation.

For most terrestrial navigation problems, the dynamics of the system are modelled using a Random Walk (RW) model. This model is widely adopted for its simplicity in determining the transition matrix due to its linear and time-invariant properties (Brown & Hwang, 1997). In this kinematic approach, the velocity states are assumed to follow a random walk process. A block diagram of the RW model is shown in Figure 6.3 and shows that the system is driven by white noise, which corresponds to the acceleration input. The dynamics matrix of the RW model based on the state

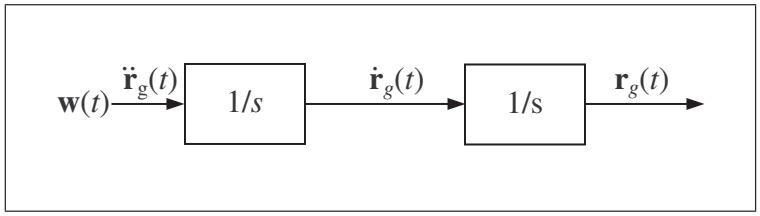


Figure 6.3: *Representation of the random walk model for velocity states.*

parameters specified is

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{M_h} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{N_h \cos\varphi} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{6.20}$$

where M_h is the meridian radius of curvature plus the height and N_h is the prime vertical radius of curvature plus the height. As shown in Equation (6.20), the form of the dynamics matrix of the RW model is quite simple. Using (6.8), the transition matrix of the RW model is derived as

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 0 & 0 & \frac{\Delta t}{M_h} & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{\Delta t}{N_h \cos\varphi} & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{6.21}$$

There is one problem associated with the transition matrix and that is its dependency

upon M_h and N_h. For formation flying, the value of these parameters can vary over time. In order for the propagation to be valid throughout time, the transition matrix must be updated at every sampling epoch. Using Equation (6.10), the process noise of the RW model is derived as

$$\mathbf{Q} = \begin{bmatrix} \frac{S_\varphi \Delta t^3}{3M_h^2} & 0 & 0 & \frac{S_\varphi \Delta t^2}{2M_h} & 0 & 0 \\ 0 & \frac{S_\lambda \Delta t^3}{3N_h^2 \cos^2\varphi} & 0 & 0 & \frac{S_\lambda \Delta t^2}{2N_h \cos\varphi} & 0 \\ 0 & 0 & \frac{S_h \Delta t^3}{3} & 0 & 0 & \frac{S_h \Delta t^2}{2} \\ \frac{S_\varphi \Delta t^2}{2M_h} & 0 & 0 & S_\varphi \Delta t & 0 & 0 \\ 0 & \frac{S_\lambda \Delta t^2}{2N_h \cos\varphi} & 0 & 0 & S_\lambda \Delta t & 0 \\ 0 & 0 & \frac{S_h \Delta t^2}{2} & 0 & 0 & S_h \Delta t \end{bmatrix}, \tag{6.22}$$

where S_ϕ, S_λ and S_h are the spectral densities of the driving white noise in north, east and up direction.

6.3.2 Reduced Dynamic Approach

Propagation in the Hill Frame

In the previous section, an orbit propagation model based on the kinematics of formation flying spacecraft was presented. In such an approach, the gravitational force was not considered. To enhance the performance of the prediction stage, a reduced dynamic approach will be demonstrated in this section.

For the reduced dynamic approach, the HE is used to account for the linearized relative central gravitational force between spacecrafts. Here, the HE is again presented but expressed in matrix form.

$$\ddot{\mathbf{r}}_h(t) + \mathbf{A}\dot{\mathbf{r}}_h(t) + \mathbf{B}\mathbf{r}_h(t) = \mathbf{w}_h(t) \quad \mathbf{Q}_h(t), \tag{6.23}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & -2n & 0 \\ 2n & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{6.24}$$

and represents the Coriolis force created from motions within the rotating triad and

$$\mathbf{B} = \begin{bmatrix} -3n^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & n^2 \end{bmatrix} \tag{6.25}$$

is a combination of the centrifugal acceleration and the linearized central gravitational force, $\mathbf{w}_h$ is the white noise driving the system, $\mathbf{Q}_h$ is the corresponding covariance matrix and n is the orbital frequency. Like any other dynamic system, the HE can be illustrated by a block diagram as shown in Figure 6.4. This is similar to the RW model; the difference is that it has feedback forces coming from its position and velocity states to capture the kinematics within a rotation frame. By using

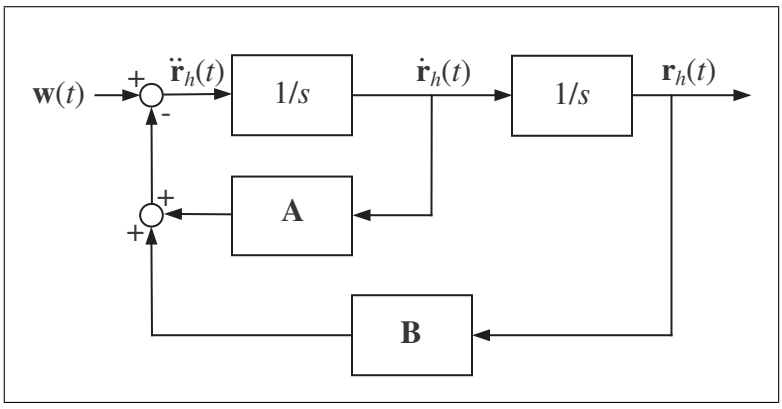


Figure 6.4: *Representation of the Hill equations of motion (Hill, 1878).*

the state space model, the 2nd order HE can be reduced to a 1st order differential

equation.

$$\dot{\mathbf{x}}_h(t) = \mathbf{F}_h\mathbf{x}_h(t) + \mathbf{G}_h\mathbf{w}_h(t) \quad \mathbf{Q}_h(t) \tag{6.26}$$

where $\mathbf{x}_h$ is the state vector containing both the position and the velocity vector in the Hill frame,

$$\mathbf{F}_h = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{B} & -\mathbf{A} \end{bmatrix} \tag{6.27}$$

is the dynamic matrix of the HE and

$$\mathbf{G}_h = \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \tag{6.28}$$

is the coefficient matrix. Solving the 1st order differential equation with Equation (6.8) leads to the following transition matrix.

$$\mathbf{\Phi}_h = \begin{bmatrix} 4-3c & 0 & 0 & s/n & -2(c-1)/n & 0 \\ 6(s-n\Delta t) & 1 & 0 & 2(c-1)/n & (4s-3n\Delta t)/n & 0 \\ 0 & 0 & c & 0 & 0 & s/n \\ 3ns & 0 & 0 & c & 2s & 0 \\ 6n(c-1) & 0 & 0 & -2s & 4c-3 & 0 \\ 0 & 0 & -ns & 0 & 0 & c \end{bmatrix}, \tag{6.29}$$

where

$$c = \cos n\Delta t, \tag{6.30a}$$

$$s = \sin n\Delta t. \tag{6.30b}$$

To obtain a simpler process noise matrix, the small angle approximation is used. Based-on Equation (2.12), the orbital frequency of a LEO geodetic mission is approximately 0.0012 rad/s. For a long sampling rate of 10 s, the motion of the anomaly

cannot exceed 1°, hence, the approximation error is less than 0.01%. Applying the small angle approximation and neglecting high order terms, the expression of the process noise matrix is

$$\mathbf{Q}_h = \begin{bmatrix} \frac{S_x\Delta t^3}{3} & \frac{S_{yx}n\Delta t^4}{4} & 0 & \frac{S_x\Delta t^2}{2} & \frac{S_{y2x}n\Delta t^3}{3} & 0 \\ \frac{S_{yx}n\Delta t^4}{4} & \frac{S_y\Delta t^3}{3} & 0 & \frac{S_{2yx}n\Delta t^3}{3} & \frac{S_y\Delta t^2}{2} & 0 \\ 0 & 0 & \frac{S_z\Delta t^3}{3} & 0 & 0 & \frac{S_z\Delta t^2}{2} \\ \frac{S_x\Delta t^2}{2} & \frac{S_{2yx}n\Delta t^3}{3} & 0 & S_x\Delta t & S_{yx}n\Delta t^2 & 0 \\ \frac{S_{y2x}n\Delta t^3}{3} & \frac{S_y\Delta t^2}{2} & 0 & S_{yx}n\Delta t^2 & S_y\Delta t & 0 \\ 0 & 0 & \frac{S_z\Delta t^2}{2} & 0 & 0 & S_z\Delta t \end{bmatrix} \tag{6.31}$$

where

$$S_{yx} = S_y - S_x, \tag{6.32a}$$

$$S_{y2x} = S_y - 2S_x, \tag{6.32b}$$

$$S_{2yx} = 2S_y - S_x, \tag{6.32c}$$

and S_x, S_y and S_z are the power spectral densities of white noise in the radial, along-track and across-track directions respectively.

Transformation to Local Geodetic Frame

Though the transition and process noise matrices for the HE are determined, they are valid only for coordinates in the Hill frame. In order to apply the dynamic model to the state parameters specified, a transformation matrix is necessary to propagate the matrices to another coordinate system where the state parameters are already defined, hence

$$\mathbf{\Phi} = \mathbf{T}^{-1}\mathbf{\Phi}_h\mathbf{T} \tag{6.33}$$

$$\mathbf{Q} \;=\; \mathbf{T}^{-1}\mathbf{Q}_h\mathbf{T} \tag{6.34}$$

The coordinate system chosen is the local geodetic frame since a local system can simplify the form of the transformation matrix when certain constraints are applied.

Before the derivation of the transformation matrix, relative positions and velocities are needed for orbit propagation with the HE. To conserve the original unit of the state vector, the relative position is expressed in a spherical measure in a local geodetic frame.

$$\Delta\mathbf{r}_l = \begin{bmatrix} \varphi_{\text{deputy}} - \varphi_{\text{chief}} \\ \lambda_{\text{deputy}} - \lambda_{\text{chief}} \\ h_{\text{deputy}} - h_{\text{chief}} \end{bmatrix} \tag{6.35}$$

As for the relative velocity, it can be determined by directly removing the chief velocity from the deputy.

$$\Delta\dot{\mathbf{r}}_l = \dot{\mathbf{r}}_{\text{deputy}} - \dot{\mathbf{r}}_{\text{chief}} \tag{6.36}$$

Although the chief and deputy velocities are in different local systems, the transformation matrix will be compensating this effect.

With the local geodetic coordinates defined, the transformation matrix can now be derived. The coordinate transformation from the local geodetic to the Hill frame is

$$\mathbf{S} = \mathbf{R}_3(u)\mathbf{R}_1(I)\mathbf{R}_3(\Omega)\mathbf{R}_3(-\text{GAST})\mathbf{R}_3(\pi-\lambda)\mathbf{R}_2(\pi/2-\varphi)\mathbf{P}_2 \tag{6.37}$$

which is composed of a series of rotation matrices derived from Section 2.1.1. If this full equation is used, then not only is the curvilinear position (φ,λ) needed, but the

GAST and Kepler elements (u,I,Ω) are also required, which would generate additional computational load to determine these parameters. To overcome this problem, the formation flying spacecraft are constrained to equatorial orbits. By doing so, the series of rotations in matrix $\mathbf{S}$ can be significantly simplified to a constant matrix independent from the positions of the spacecraft.

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{6.38}$$

With regard to the transformation of the velocity, when using the product rule, transformation dependency upon both the position and velocity states is established. The assumption made here is that the variation is only within the anomaly and the rotation of the Earth. All other disturbing motions due to higher order gravitations are neglected. Applying these conditions to $\mathbf{S}$ and taking the time derivative gives

$$\dot{\mathbf{S}} = (n - \omega_e) \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \tag{6.39}$$

where ω_e is the rotational rate of the Earth. The final form of the transformation matrix from the local geodetic to the Hill frame is

$$\mathbf{T} = \begin{bmatrix} \mathbf{S} & \mathbf{0} \\ \dot{\mathbf{S}} & \mathbf{S} \end{bmatrix}. \tag{6.40}$$

Many of the matrices derived in this section have reliance on the Hill frame. As shown in Chapter 3, the proper selection of the Hill frame is important. The ideal location of the Hill frame is in the centre of the configuration. This is achievable by

the KF since the Hill frame can be updated at every sampling epoch. Given such a condition, the semi-major axis of the Hill frame can be determined from calculating the average height of the two spacecraft.

$$a_h = \frac{h_{\mathbf{deputy}} + h_{\mathbf{chief}}}{2} + a_e \tag{6.41}$$

To verify the accuracy of the transformation matrix, it should be compared with direct linear mapping developed by Alfriend & Schaub (2000). This analysis is performed on the 2-by-1 elliptical configurations with 0.5, 3 and 10 km separations. The average altitude of the spacecraft is approximately 475 km above the Earth's surface and the orbit is simulated with a J_2 perturbation. The errors committed by the transformation matrix are shown in Figures 6.5, 6.6 and 6.7. Note that these are not the errors expected from the navigation solution of the EKF. These errors represent only the misinterpretation of the HE for orbit propagation. The plots indicate that the errors appear only in the radial and along-track directions and the levels of these errors are directly proportional to the separation distances and purely periodic. The transformation is biased for the radial position and along-track velocity. There are no transformation error in the across-track direction. This is expected as the motion of the satellites are concentrated in the elliptical plane. For the elliptical configuration with a 0.5 km separation, the transformation error is insignificant at the millimetre level. However, when the separation distance is extended to 10 km, errors can reach 3 m and 3 cm/s. This analysis shows that the derived transformation matrix is only an approximation. When using the reduced dynamic approach for orbit propagation, degradation in the navigation performance is expected as the separation distances are extended.

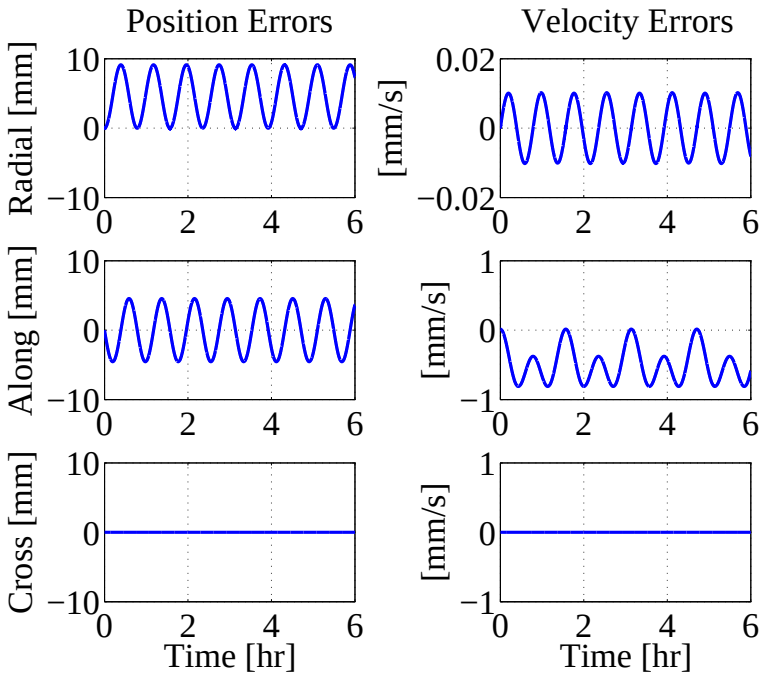


Figure 6.5: *Transformation errors for a 0.5 km elliptical configuration.*

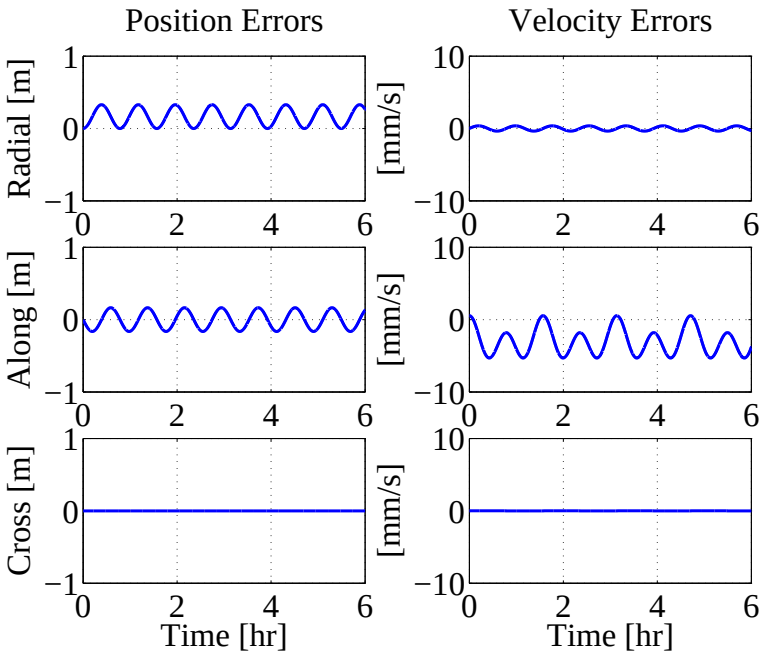


Figure 6.6: *Transformation errors for a 3 km elliptical configuration.*

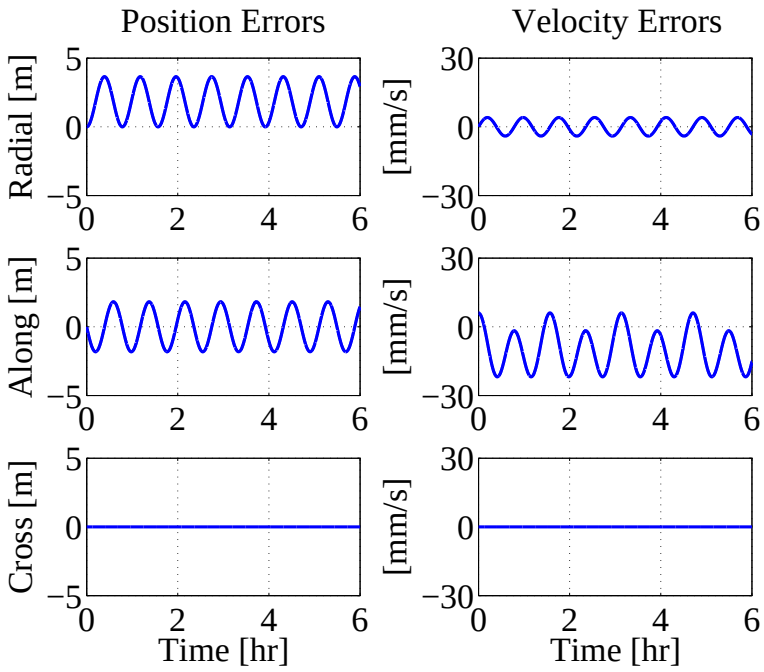


Figure 6.7: *Transformation errors for a 10 km elliptical configuration.*

6.4 Ambiguity Resolution

The Extend Kalman Filter (EFK) itself is a float filter, therefore it does not take the integer nature of ambiguities into account. In order to exploit the full potential of the carrier phase measurements, an additional procedure is needed for estimation of integer ambiguities. Over the years, various methods for resolution of Double Difference (DD) integer ambiguities have been developed and implemented. These methods include,

- Rounding
- Ambiguity function method (Counselman, 1981)
- Fast Ambiguity Resolution Approach (FARA) (Frei & Beutler, 1990)

- Least squares ambiguity search technique (Hatch, 1991)
- Fast Ambiguity Search Filter (FASF) (Chen & Lachapelle, 1994)
- Least squares AMBiguity Decorrelation Adjustment (LAMBDA) (Teunissen, 1995)

Among these methods, the Least squares AMBiguity Decorrelation Adjustment (LAMBDA) is consider the most optimal. Teunissen (1998) has proven that the success rate of LAMBDA is always greater than or equal to any other integer ambiguity estimator. Due to this reason it has been widely adapted and will be selected for this research. In addition to the estimation procedure, a validation scheme is also needed to verify the reliability of the estimated ambiguities. Since any erroneous integers can cause the filter to diverge over time and degrade the relative positioning solution significantly. The algorithms for ambiguity estimation and validation are discussed in the following subsections.

6.4.1 Ambiguity Estimation

Due to the differential process between satellites, DD ambiguities are highly correlated among themselves. In the ambiguity domain, the search space is a highly elongated ellipsoid. To address such a problem, Teunissen (1995) has proposed the Least squares AMBiguity Decorrelation Adjustment (LAMBDA) to reduce the number of possible ambiguity sets to be searched. With the search space reduced, the objective is to find the integer set N that has the minimum sum of squared residuals (least-

squares). To simplify the representation, the DD operator $\Delta\nabla$ is dropped.

$$(\mathbf{N} - \hat{\mathbf{N}})^{\mathsf{T}} \mathbf{Q}_{\hat{\mathbf{N}}}^{-1} (\mathbf{N} - \hat{\mathbf{N}}) = \min \tag{6.42}$$

where $\mathbf{N}$ is the true value of the integer ambiguity set, $\hat{\mathbf{N}}$ are the float estimates of the ambiguities provided by the Kalman filter and $\mathbf{Q}_{\hat{\mathbf{N}}}$ associated covariance matrix for the ambiguities. The constraint defined in Equation (6.42) also represents the ambiguity search region.

$$(\mathbf{N} - \hat{\mathbf{N}})^{\mathsf{T}} \mathbf{Q}_{\hat{\mathbf{N}}}^{-1} (\mathbf{N} - \hat{\mathbf{N}}) \leq \chi^2 \tag{6.43}$$

where χ^2 is a positive constant that controls the size of the ellipsoidal search space. The shape of the ellipsoid is governed by the covariance matrix and centered at the float solution. For DD ambiguities, where high correlations exist, the ellipsoid is highly elongated. To achieve better search efficiency, LAMBDA is used to transform the ambiguity set to another domain where the covariance matrix becomes nearly diagonal. The transformation is referred as the Z-transformation and is defined as follows.

$$\hat{\mathbf{M}} = \mathbf{Z}\hat{\mathbf{N}} \tag{6.44a}$$

$$\mathbf{Q}_{\hat{\mathbf{M}}} = \mathbf{Z}^{\mathsf{T}} \mathbf{Q}_{\hat{\mathbf{N}}} \mathbf{Z} \tag{6.44b}$$

where $\hat{\mathbf{M}}$ is the transformed float ambiguity set in the Z-domain, $\mathbf{Z}$ is the transformation matrix and $\mathbf{Q}_{\hat{\mathbf{M}}}$ is the decorrelated covariance matrix. Due to the integer nature of the transformation matrix $\mathbf{Z}$, the resulting decorrelated covariance matrix $\mathbf{Q}_{\hat{\mathbf{M}}}$ is not totally diagonal. The covariance parameters are however significantly reduced which creates a much more uniform ellipsoid search space. Performing the ambiguity

search in the z-domain, then the sequential bounds for the individual ambiguities are (Wu, 2003),

$$\begin{aligned}
(M_1 - \hat{M}_1)^2 &\leq \sigma^2_{\hat{M}_1}\chi^2, \\
(M_{2|1} - \hat{M}_2)^2 &\leq \sigma^2_{\hat{M}_{2|1}}\left(\chi^2 - \frac{\left(M_1 - \hat{M}_1\right)^2}{\sigma^2_{\hat{M}_1}}\right), \\
&\vdots \\
(M_{n|N} - \hat{M}_n)^2 &\leq \sigma^2_{\hat{M}_{n|N}}\left(\chi^2 - \sum_{i=1}^{n-1}\frac{\left(M_i - \hat{M}_{i|I}\right)^2}{\sigma^2_{\hat{M}_{i|I}}}\right),
\end{aligned} \tag{6.45}$$

where $M_{n|N}$ are the conditional ambiguities, given by

$$\begin{aligned}
M_{2|1} &= M_2 - \frac{\sigma_{\hat{M}_2\hat{M}_1}\left(M_1 - \hat{M}_1\right)}{\sigma^2_{\hat{M}_1}}, \\
&\vdots \\
M_{n|N} &= M_n - \sum_{i=1}^{n-1}\frac{\sigma_{\hat{M}_n\hat{M}_{i|I}}\left(M_{i|I} - \hat{M}_n\right)}{\sigma^2_{\hat{M}_i|I}},
\end{aligned} \tag{6.46}$$

and $\sigma_{\hat{M}_{n|N}}$ are the conditional variances

$$\sigma^2_{\hat{M}_{n|N}} = \sigma^2_{\hat{M}_n} - \sum_{i=1}^{n-1}\frac{\sigma^2_{\hat{M}_n\hat{M}_{i|I}}}{\sigma^2_{\hat{M}_{i|I}}}. \tag{6.47}$$

Throughout the search procedure, two successful ambiguity candidates with the first and second lowest sum of squared residuals are kept. The two ambiguity sets will be used to validate the reliability of the best estimated ambiguities.

6.4.2 Ambiguity Validation

Before the DD integer ambiguity estimates are selected for the navigation solution, a validation process is needed to assess the correctness of the selected ambiguities,

since any erroneous integer values can significantly degrade the positioning accuracy and over time causing the filter to diverge. To validate the ambiguity set estimated, a reliability test is conduct to determine the level of confidence on the solution provided by LAMBDA. The reliability test is carried out using a ratio test, where the best ambiguity solution set is tested against the second best solution. Here the best is determined through the sum of squared residuals Ω.

$$\frac{\Omega_2}{\Omega_1} > \delta \tag{6.48}$$

where Ω_1 and Ω_2 represents the sum of squared residuals of the best and second best ambiguity sets, respectively, and δ is the critical value for the ratio test. It is very common to assume that the test statistic follows the F-distribution. In such a case, the critical value can be determined based on the significance level and number of ambiguities (degrees of freedom). However, according to Lu (1995) this assumption is not correct, since the ambiguity sets are not independent, but actually correlated. Though using the F-distribution for the critical value may not be correct statistically, but it is often used and seems to work satisfactorily (Frei & Beutler, 1990; Abidin, 1993). For this research, the determination of the critical value is based on the F-distribution.

For the null hypothesis H_0, the test statistic is greater than the threshold provided by the F-distribution and accepts the best ambiguity set as the correct solution. As for the alternative hypothesis H_1, the test statistic is lower than the threshold and it rejects the best solution sets. In such a test, the best integer solution set is only accepted when it can be discriminated significantly from any other solution sets. If the test passes for the second best solution then it will pass for any other ambiguity

set, since the sum of squared residuals will always be larger.

Committing a type-I error can occur during reliability testing. Here, the null hypothesis is rejected while it is actually true, which causes the search procedure to take a longer time to resolve integers. As for the type-II error, the null hypothesis is accepted while it is actually false. In such a case, the incorrect ambiguity set is accepted, which can cause the filter to diverge over time. Since its more crucial to avoid committing a type-II error, a significant level of 10% is chosen for the F-distribution. Such a significant level may lengthen the resolution time, but is better than choosing the wrong ambiguity set.

Chapter 7

Test Scenarios & Results

In this chapter, the navigation algorithms that were presented in the previous chapter are evaluated by different simulated test scenarios. The details of each scenario are thoroughly described and followed by the corresponding results. The results are then used to address the relative navigation performance in terms of the positions and velocities accuracies, the ambiguity resolution performance and the robustness during signal outages.

7.1 Test Scenarios

7.1.1 Orbit Configuration

As an alternative to analyzing the simple leader-follower formations whereby the baseline distance remains fairly consistent throughout the orbit, the 2-by-1 elliptical formation described in Section 3.2.2 are examined. To simplify the simulation procedure, the arrangement of all formations involves only two spacecraft. Furthermore, in order to maintain symmetry within the configuration and at the same time maximize the separation distances between spacecraft for analysis purposes, all formations are symmetrically structured. With the formation style defined, different configuration sizes are used to assess the performance over different separation distances. Configurations with 0.5, 3 and 10 km separations are chosen for assessment because these

short and medium baselines can fulfil the requirement of the linearized HE. Choosing extended separation distances would cause the HE model to fail because the relative dynamics can no longer be approximated by a linear model. The absolute orbit is defined by the general mean orbital elements shown in Table 7.1, where b is the separation distance of the configuration. All spacecraft that are simulated are placed in

Table 7.1: *Mean orbit elements of spacecraft in 2-by-1 elliptical formations.*

Initial Mean Orbit Elements	
a	a_e + 485 km
e	$b/4a$
I	0°
Ω	0°
ω	0°, 180°
ν	180°, 0°
date	August 1, 2004
time	Midnight

equatorial orbits flying at a mean altitude of 485 km. Spacecraft at this low altitude are travelling at an extremely high velocity of 7.1 km/s. As for the eccentricity, it has been illustrated in Chapter 3 that it is one of the key elements for controlling the size of the elliptical motion. Thus, the eccentricities of the spacecraft are separation dependent, as shown in Table 7.1. The perigee and the true anomaly of the spacecraft within the same formation are differentiated by 180° in order to achieve the symmetric formation. All spacecraft are initiated at the Greenwich meridian during local midnight but varying in altitude.

7.1.2 GPS Errors

With the formations defined, the next procedure is to simulate the GPS errors that are potentially experienced by the spacecraft. The major error sources affecting the GPS measurements have been discussed thoroughly in Chapter 4 and so it is only mentioned briefly in this section. To simulate the GPS measurements and any possible errors, the modified version of SIMGNSSII^{TM} described in Chapter 5 is used.

One of the main concerns with GPS navigation is the effect that ionosphere has on the pseudorange and carrier phase measurements. Knowing that the variation of the ionosphere can be quite significant, two ionospheric conditions, low and moderate, are chosen for testing the navigation algorithms. The relative error level in low ionospheric condition is approximately 2 PPM and 7 PPM in moderate condition. A visualization of code delays experienced by the spacecraft in both the low and moderate ionospheric conditions are shown in Figure 7.1. As the figure shows, the

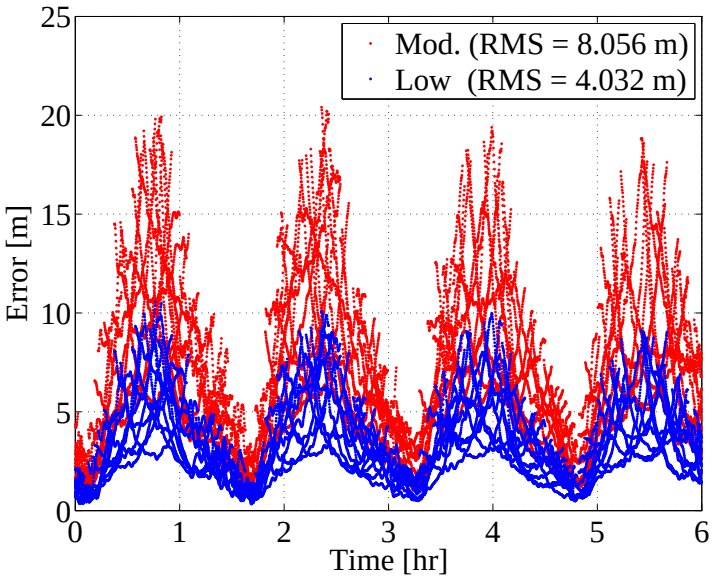


Figure 7.1: *Code delays in low and moderate ionospheric conditions.*

magnitude of the ionospheric error in moderate conditions is twice the magnitude of those in the low ionospheric condition. It is possible for the absolute range errors to exceed 20 m during ionospheric peaks. Since the spacecraft have orbital periods of 1.5 hours, four revolutions around Earth are made over the six hour simulation time, which explains the four ionospheric peaks that are shown.

Other simulated errors include multipath errors and receiver noise errors. These errors are assumed to be uncorrelated between receivers, therefore their relative error levels can be calculated directly from the absolute error levels. Table 7.2 is a summary of the error levels for each of the sources. Note that the L_1 Doppler

Table 7.2: RMS *level of* GPS *errors.*

Orbital [PPM]	Ionosphere [PPM]		Multipath [mm]			Receiver Noise [mm]/[mm/s]			
	Low	Mod.	p_{CA}	ϕ_{L1}	ϕ_{L2}	p_{CA}	ϕ_{L1}	ϕ_{L2}	$\dot{\phi}_{L1}$
0.1	2	7	218	5.8	7.3	50	0.8	1.0	15

measurement is only contaminated by 1.5 cm/s of Gaussian noise. Effects from the ionospheric and multipath errors have been neglected, therefore the observation is relatively optimistic.

The GPS satellite availability and corresponding Dilution Of Precision (DOP) values are shown in Figure 7.2. The GPS antennas are assumed to point in zenith direction at all times. The cutoff angle used was 5° and the availability varies between 7 to 13 satellites. This low cutoff angle is achievable because there are usually no obstructions at low elevation in space. Note that the number of GPS satellites visible changes quite rapidly over time. This is due to the high velocity of the LEO spacecraft. Overall the satellite availability is acceptable, therefore, low DOP values are common

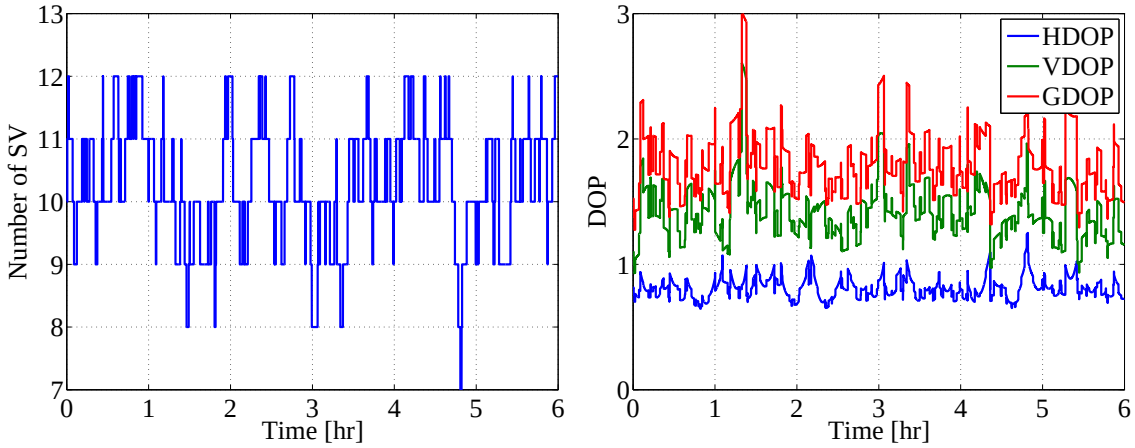


Figure 7.2: GPS *satellite availability and dilution of precision for* LEO *spacecraft.*

throughout.

The combination of two ionospheric conditions and three formations discussed in the previous section means that a total of six test scenarios, summarized in Table 7.3, are used to assess the navigation performance.

Table 7.3: *Description of test scenarios.*

Test	Baseline	Ionospheric Condition
1	0.5 km	low
2	0.5 km	moderate
3	3 km	low
4	3 km	moderate
5	10 km	low
6	10 km	moderate

7.1.3 Processing Procedure

The general simulation and processing procedures are summarized in Figure 7.3. The trajectories of the spacecraft are generated with an orbit integrator written

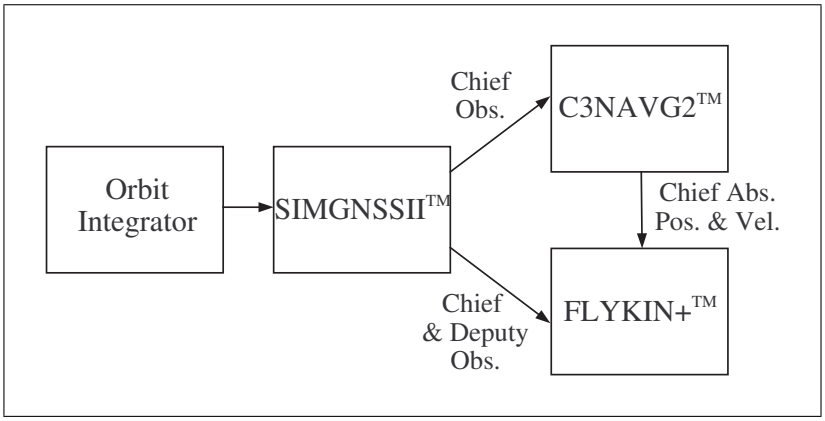


Figure 7.3: *Simulation and processing procedures.*

in MATLAB®. The only perturbation that are included in the integration is the J_2 disturbing potential, discussed in Section 2.3. All higher order perturbations, atmospheric drag and solar pressure are neglected. This is followed by the simulation of GPS measurements and errors with SIMGNSSII™, which has been thoroughly discussed in Chapter 5. The simulated measurements of the reference spacecraft are then processed by C³NAVG²™ to derive the absolute position and velocity using only pseudoranges and L_1 Doppler measurements. Examples of the absolute position and velocity errors can be found in Appendix, Figure A.1. Finally, the absolute trajectory is fed into FLYKIN+™ to estimate the relative position and velocity of the spacecraft based-on precise carrier phase measurements which requires an ambiguity resolution process.

7.1.4 Sensitivity Analysis

To obtain optimal navigation solutions for the spacecraft, a sensitivity analysis is conducted in order to determine the actual process noise for each model and forma-

tion. This procedure is necessary because according to Gelb et al. (1974), optimal estimation can only be achieved when correct statistical information are used. Such analysis also illustrates the importance of having the correct process noise. The analysis is carried out by comparing the navigation performance in terms of position and velocity estimates over a range of Spectral Densities (SPD). In this experiment only L_1 and WL measurements are used.

The analysis is first conducted for the elliptical formations with a distance of 0.5 km. Figure 7.4 shows the position and velocity errors in radial, along-track and cross-track for both the RW model and the HE. The horizontal axes are the design

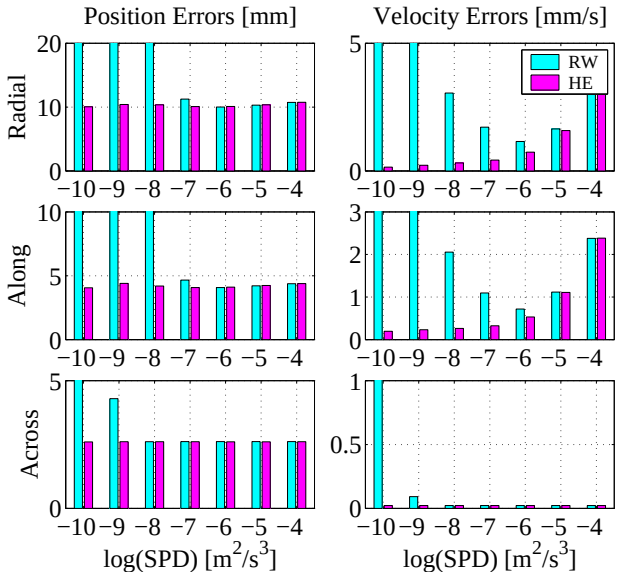


Figure 7.4: *Sensitivity analysis for elliptical formation with 0.5 km separation.*

parameters, which represent the logarithm of the SPD. The magenta error bars correspond to the approach using the RW model and the cyan bars represent the errors for HE. Due to the nature of GPS, the largest errors are found in the radial direction as satellites are only available above the horizon. As for the accuracy in the

horizontal plane, the across-track component is slightly better than the along-track. This is because there is zero motion in the across-track direction, which makes it easier to model.

Based on the analysis of the errors associated with the RW model, the trend of the errors is parabolic. SPD values that are too large or too small cause the error levels to significantly increase. This indicates that the optimal SPD values must be located at the minimum of the curve. Based on this assumption, the optimal SPD chosen for the RW model in the radial and along-track direction is $10^{-6}\,\mathrm{m}^2/\mathrm{s}^3$. However, the across-track component is not as straight-forward and requires more investigation. Knowing that there is zero motion in the across-track direction, it seems straight-forward that the RW model would be able to model its behaviour easily and the error level would have not significantly increased even when the SPD is overly suppressed. However, this is contradictory with the across-track error bars shown in Figure 7.4.

To investigate the sensitivity of the across-track error, the relative position errors for a SPD of $10^{-9}\,\mathrm{m}^2/\mathrm{s}^3$ are plotted in Figure 7.5. The plots indicate that the

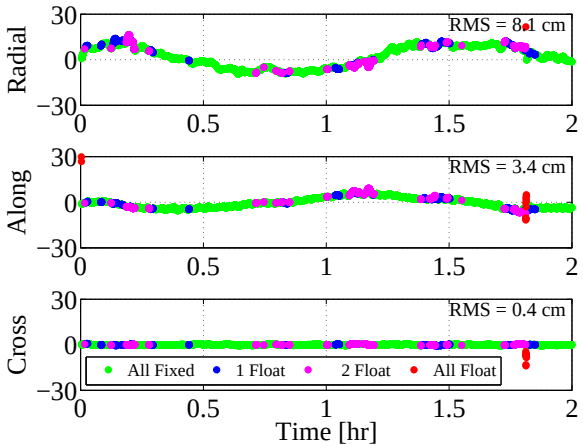


Figure 7.5: *Relative position errors for* RW *model with* SPD *of* $10^{-9}\,\mathrm{m}^2/\mathrm{s}^3$.

across-track error is quite consistent throughout, except during the period where the ambiguity resolution failed (red). The failure with the ambiguity resolution is related to the high radial and along-track errors, which are the results of over-suppressing the SPD in the associated directions and not related to the across-track direction itself. The strong period trends are evidence of over-constraining the SPD level in the radial and along-track directions, which is not visible in the across-track error. Since the increase of errors is not related to the across-track SPD, a much lower SPD magnitude of $10^{-10}\,\mathrm{m}^2/\mathrm{s}^3$ is chosen for the across-track direction.

Analyzing the errors for the HE model, Figure 7.4 shows that they are very similar to the RW model when SPD levels are high (10^{-4} to $10^{-5}\,\mathrm{m}^2/\mathrm{s}^3$), where the navigation solution is weighed more heavily towards the GPS measurements. As the SPD is decreased and more weight is shifted onto the process model, it is evident that the HE demonstrated better navigation performance. The position errors are very consistent and deviate little for the range of the SPD tested. As for the velocity errors, the trend followed an exponential curve where insignificant improvements are achieved for SPD lower than $10^{-8}\,\mathrm{m}^2/\mathrm{s}^3$. Based on such an analysis, the SPD chosen for the HE are 10^{-8}, 10^{-8} and $10^{-10}\,\mathrm{m}^2/\mathrm{s}^3$ for the radial, along-track and cross-track directions respectively. The sensitivity analysis is repeated for the formations with 3 and 10 km separations. Their sensitivity analysis plots are shown in Figure A.2(a) and A.2(b).

Based on the sensitivity analysis, the optimal SPDs are listed in Table 7.4. As already mentioned, there is zero cross-track motion in the elliptical formation, therefore the SPD for this component is extremely small ($10^{-10}\,\mathrm{m}^2/\mathrm{s}^3$) and also the same for all configuration sizes. As for the radial and along-track SPDs, their levels of

Table 7.4: *Logarithm of the optimal* SPD *for elliptical formations.*

	0.5 km		3 km		10 km	
	RW	HE	RW	HE	RW	HE
Radial	-6	-8	-5	-7	-5	-7
Along	-6	-8	-5	-7	-5	-7
Across	-10	-10	-10	-10	-10	-10

motion are similar and therefore have the same magnitude. Furthermore, Table 7.4 shows that the SPD levels increase in proportion to the configuration size, which is due to the increase in relative dynamics between the spacecraft. Although no extensive analysis has yet been done to compare the two propagation models used, the sensitivity analysis conducted does indicate a slightly better performance by the HE. This is evident from Table 7.4, where the SPD levels of the HE are two magnitudes smaller than the RW model for all the formation sizes.

7.1.5 Performances Comparison

In GPS navigation, various types of observations can be used to derive the navigation solution. To investigate the effects of different measurement combinations with the process models, a performance comparison analysis is conducted by comparing the navigation solutions of different processing strategies. This type of analysis cannot only display the optimal measurement combinations but it can also assess the contribution of each measurement to the navigation solution. The comparison is conducted only for Test Scenario 3, where the separation of the elliptical formation is 3 km and ionospheric condition is low. The results of other test scenarios are not shown because they provide similar conclusions. The different combinations of processing strategies used are shown in Table 7.5.

Table 7.5: *Description of processing strategies.*

Strategy	Model	Observation
1	RW	p_{CA}
2	HE	p_{CA}
3	RW	p_{CA}, $\dot{\phi}_{\mathrm{L1}}$
4	HE	p_{CA}, $\dot{\phi}_{\mathrm{L1}}$
5	RW	ϕ_{L1}
6	HE	ϕ_{L1}
7	RW	ϕ_{L1}, $\dot{\phi}_{\mathrm{L1}}$
8	HE	ϕ_{L1}, $\dot{\phi}_{\mathrm{L1}}$
9	RW	ϕ_{L1}, ϕ_{WL}
10	HE	ϕ_{L1}, ϕ_{WL}

Results

The relative position and velocity errors for the different processing strategies are shown in Figure 7.6. The plots are the 3D RMS errors for each of the strategies. Starting with Strategy 1 on the left and ending with Strategy 10 on the right.

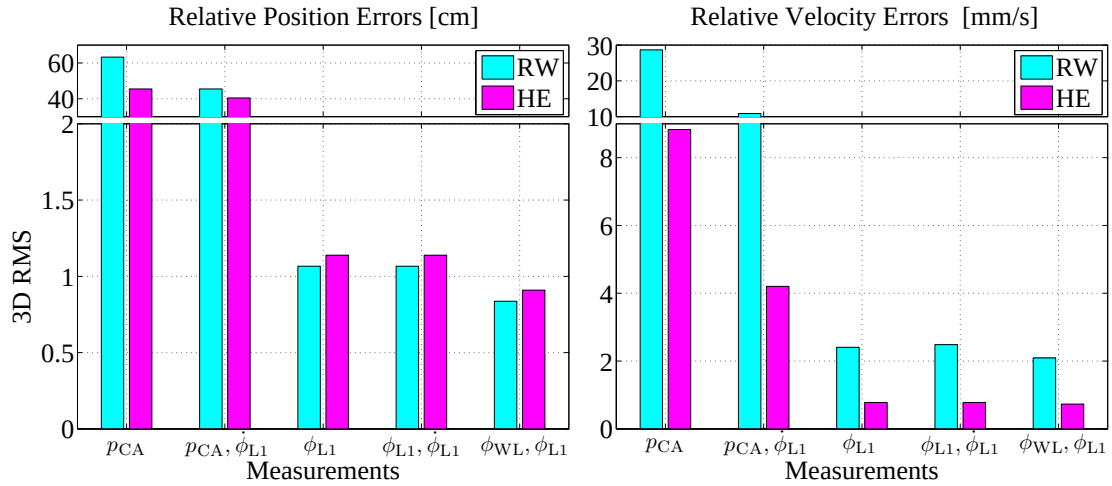


Figure 7.6: 3D RMS *errors for various processing strategies.*

Focusing on code-based relative navigation (Strategies 1 to 4), it appears that the RW model can only provide decimetre level relative position accuracy. The po-

sitioning performance can be slightly improved by the inclusion of L_1 Doppler measurements. These additional observations also improved the accuracy in the velocity domain. As for the approach with the HE, improvements are evident in the both position and velocity domains compared to the RW approach. In the case where L_1 Doppler measurements are used, relative velocity accuracy in the millimetre per second level is demonstrated.

Figure 7.6 shows that centimetre positioning accuracy can be achieved only by using carrier phase information with proper ambiguity resolution imposed (Strategies 5 to 10). Furthermore, only the carrier phase measurements themselves can provide millimetre per second velocity accuracy. The use of Doppler measurements have nearly no impact on the navigation performance for carrier-based navigation. This is because precise measurements processed at a high data rate (1 Hz) can replace the use of Doppler measurements to facilitate velocity estimation. The advantage of using the HE as the process model is the significant improvement in the velocity domain. Relative velocity errors are marginally above the 1 mm/s level when the HE is used.

The best relative navigation performance is achieved with the use of WL and L_1 carrier phase measurements (Strategies 9 & 10). Figure 7.7 is an example of the relative position and velocity errors obtained using WL and L_1 carrier phase measurements with the HE model. Results for all other strategies can be found in Appendix. The subplots correspond to a projection of the position and velocity errors onto the radial, along-track and cross-track directions. The points plotted in red represent periods when all ambiguities are unfixed, which usually occur at the beginning when all integer ambiguities are unknown. Green represents complete ambiguity resolu-

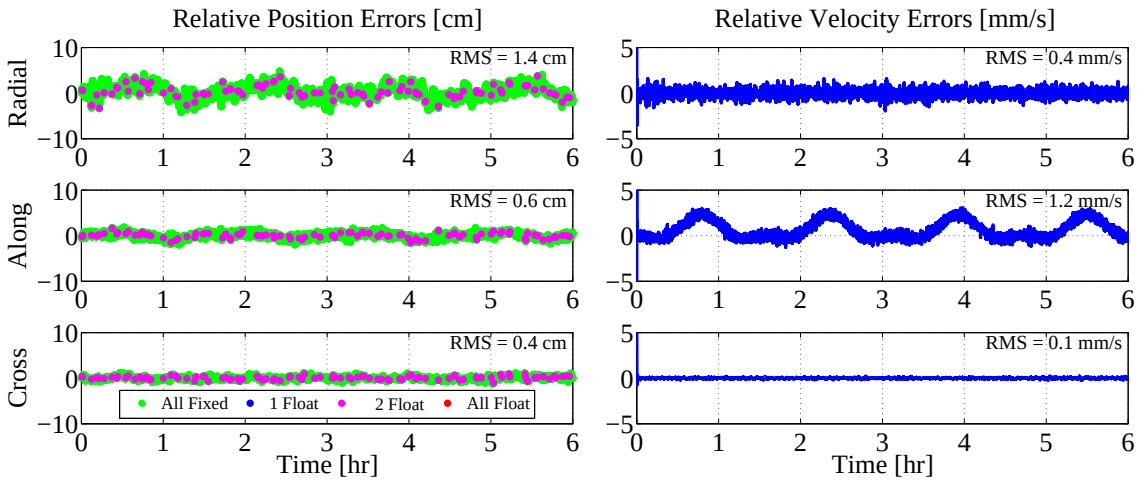


Figure 7.7: *Strategy 10 relative position and velocity errors.*

tion, where all integer ambiguities are fixed to integer values. Blue represents periods when there is one unfixed ambiguity and magenta represents the time when there are two or more unfixed ambiguities. This usually happens when new GPS satellites come into view and new ambiguities are introduced to the estimation process. Although not visible from the plots, but before ambiguity resolution, the position errors are at the metre level. Only after 11 s when all ambiguities are resolved, are the error levels significantly reduced. With ambiguities resolved, the position and velocity accuracies achieved are in the centimetre and millimetre per second level, respectively. The navigation performance is better than all other measurement combinations used. Furthermore, using the WL and L_1 also deliver the shortest fix time for integer ambiguity resolution.

7.2 Test Results

In this section the relative navigation performance for each test scenarios is investigated. The accuracy in the position and velocity domains is first assessed by analyzing the performance in a six hour period. Cold and warm start analyses are followed to investigate the performance in the ambiguity domain. All solutions presented in this section are processed with L_1 and WL carrier phase measurements to deliver the best relative navigation performance.

7.2.1 Performance Analysis

Test 1 Results

Test Scenario 1 corresponds to the elliptical formation with 0.5 km separation experiencing low ionospheric effect. Figure 7.8 displays the relative position errors (left) and velocity errors (right) for the radial (top), along-track (middle) and cross-track (bottom) directions. The plots show results obtained from both the RW (blue) and HE (red) approaches. With such a short separation distance, the differential process nearly removes all correlated errors, therefore no ionospheric effect can be seen from the error plots. The time requirement for ambiguity first fix is only 11 s for both RW and HE models. After initial fixing, all integer ambiguities are resolved throughout the course of the simulation. Comparing the results between the RW and HE approaches, slightly better smoothing in the position domain is achieved by the HE model. However, this improvement is so minor that it is not reflected in the position errors RMS. As for the velocity accuracy, significant improvement is gained by using the HE as process model. Sub-millimetre per second precision can be achieved with

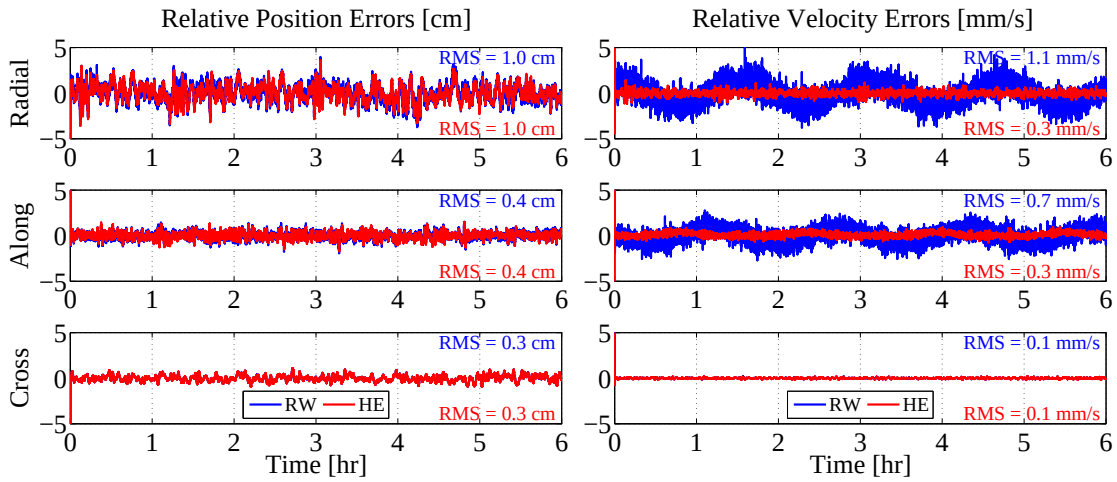


Figure 7.8: *Test 1 relative position and velocity errors.*

the HE model, where the RW approach can only provide accuracy at the millimetre per second level. This improvement is more evident in the radial and along-track directions, where the relative motion of the spacecraft are concentrated.

Test 2 Results

For Test Scenario 2, the formation separation is kept at 0.5 km but the ionospheric condition is enhanced to a moderate level. The relative position and velocity accuracies achieved are shown in Figure 7.9. Compared to the results obtained from Test 1, the performance is nearly the same and statistically there is no difference between the relative position and velocity errors. The performance in the ambiguity domain is also identical. All integer ambiguities are resolved after the initial resolution of 11 s. The results indicate that when the formation is only separated by 0.5 km, the differential process can nearly eliminate the ionospheric effect. Even when the ionospheric conditions are enhanced marginally, the change in the relative ionospheric error is still insignificant. The remaining error sources affecting the accuracy are

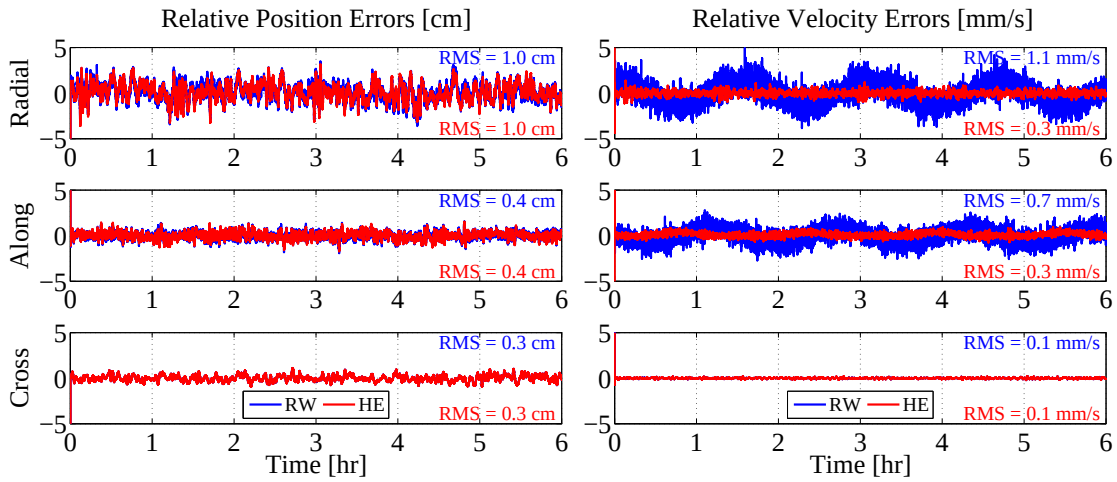


Figure 7.9: *Test 2 relative position and velocity errors.*

dominated by multipath errors and receiver noise.

Test 3 Results

In Test 3, the separation of the elliptical formation is extended to 3 km and low ionospheric conditions are simulated. The test results are shown in Figure 7.10. The initial resolution of integer ambiguities begins at 11 s for both the RW and HE

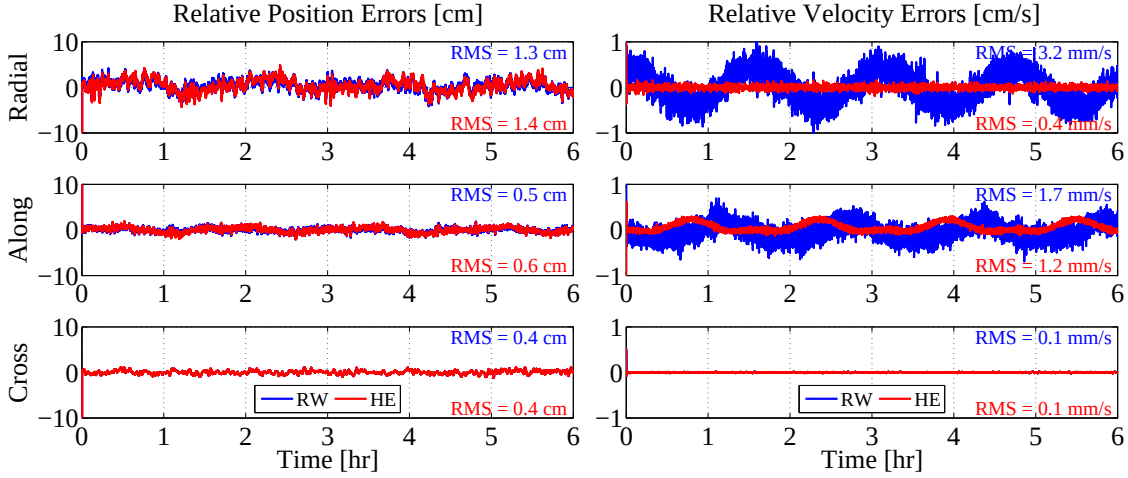


Figure 7.10: *Test 3 relative position and velocity errors.*

approaches. Their ambiguity resolution performance are nearly the same throughout the simulation. After initial resolution, 99.95% of the ambiguities are successfully resolved. The insignificant amount of unresolved ambiguities are related to new incoming measurements from GPS satellites coming into view. Most of these new integer ambiguities are resolved after one epoch. Comparing the positioning errors between the RW and HE approaches, their difference is not significant. However, the statistics does indicate that the RW model is performing slightly better in the radial and along-track direction. As for the velocity performance, again better accuracy is achieved with the HE approach by including the relative dynamics into the navigation model. Millimetre per second velocity accuracy is demonstrated by using the HE, where only few millimetres per second accuracy is obtained by the RW approach.

Test 4 Results

For Test Scenario 4, the same formation size of 3 km is used but the analysis is conducted in moderate ionospheric conditions. Figure 7.11 gives the relative navigation results in the position and velocity domains. Comparing the results to Test Scenario 3 shows the degraded navigation performance due to the amplified ionosphere. Traces of amplified ionosphere peaks can be found in the relative position errors. Analyzing the results, a slightly better positioning accuracy is again achieved by the RW model. However, HE is still delivering much better velocity estimates. Even when the ionosphere is increased to a moderate level, the same velocity accuracy is obtained by the HE approach. As for the performance in the ambiguity domain, a similar performance is obtained from both approaches. The initial ambiguity fix time is 16 s and 95.25% of the integer ambiguities are resolved. The slightly degraded

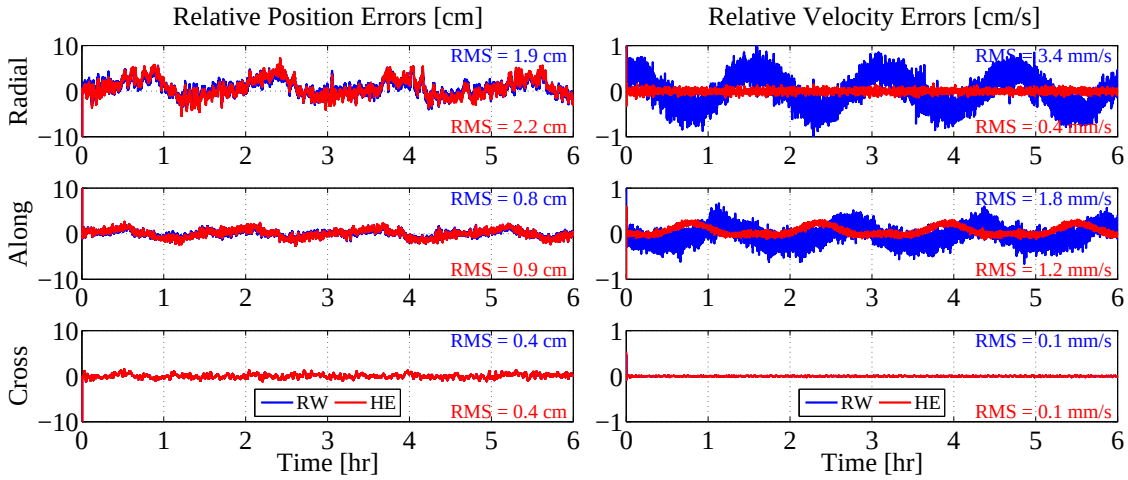


Figure 7.11: *Test 4 relative position and velocity errors.*

percentage compared to Test 3 is due to the extended resolution time for new ambiguities during ionospheric peak periods. This is reflected by the fixed ambiguity percentages shown in Figure 7.12. The bar graph shows that the percentage of fixed

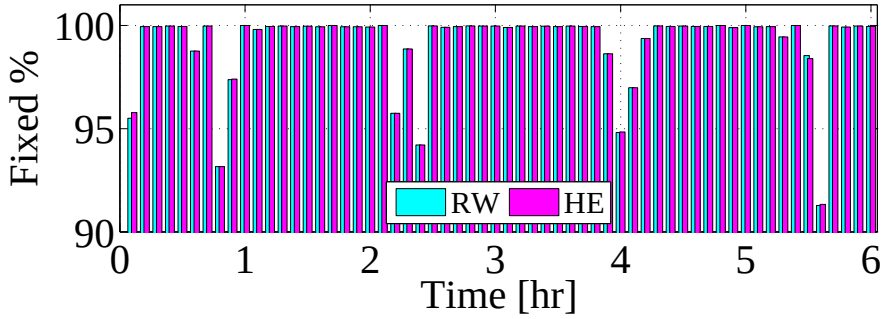


Figure 7.12: *Test 4 fixed ambiguity percentages.*

ambiguities always drops during the ionosphere peak periods, which reflects how more time is required for new ambiguity resolution during these periods.

Test 5

In Test 5, the size of the spacecraft separation is extended to a medium distance of 10 km. Ionospheric conditions are kept at low levels for this scenario. Figure 7.13 shows plots of the relative navigation errors achieved. The performance of the RW

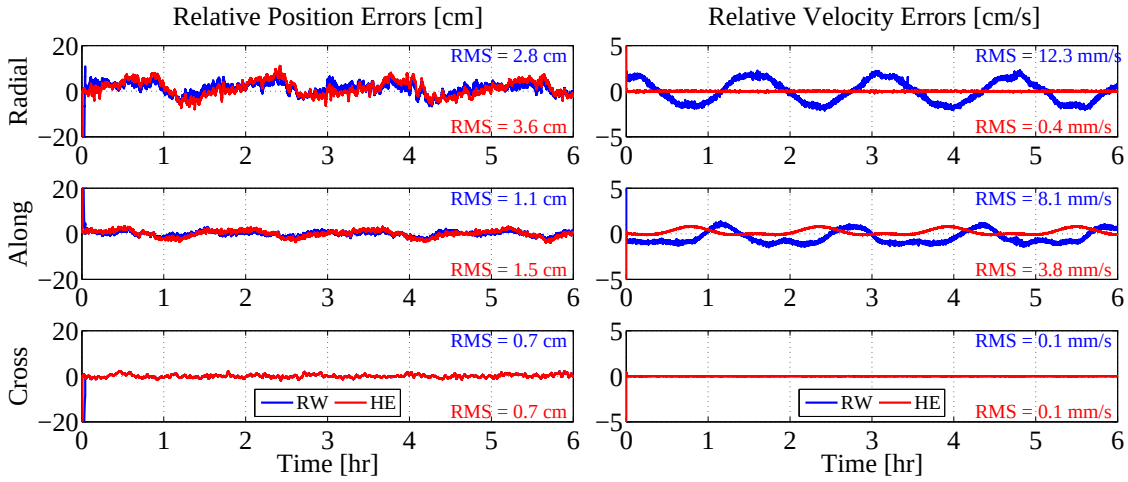


Figure 7.13: *Test 5 relative position and velocity errors.*

model in the position domain clearly outperforms the HE approach, especially in the radial direction where its is evident that the 3D RMS error is smaller by almost 1 cm. As for the velocity errors, the HE is still demonstrating better performance than the RW model. However, strong systematic errors in the along-track direction are definitely evident with the HE approach. This is related to the linearity of the HE, where the approximation is degraded due to the extended separation distance between the spacecraft.

Figure 7.14 shows the percentage of the ambiguities resolved throughout the simulation. The average percentages of fixed ambiguities are 97.27% and 97.36% for the RW and HE models respectively, which indicates that ambiguity resolution performances between the two approaches are similar after the first fix. The only

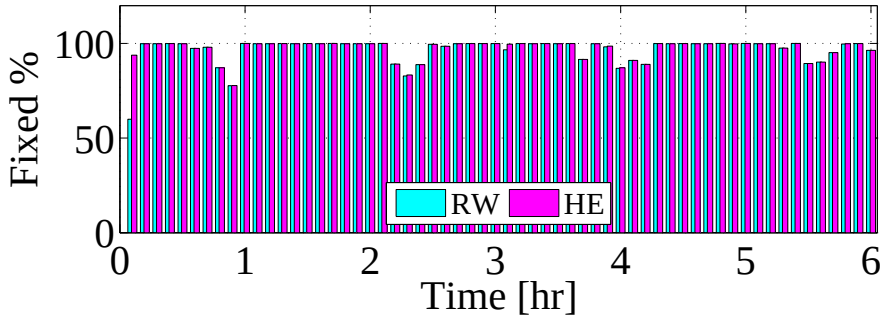


Figure 7.14: *Test 5 fixed ambiguity percentages.*

major difference is found in the initial columns, where a large discrepancy is found in the ambiguity fix time between the two approaches. For the RW model the fix time is 2 m24 s, and 22 s for the HE model. This is unexpected as similar ambiguity resolution performance has been demonstrated by the two approaches in previous Test Scenarios.

Test 6

Test scenario 6 corresponds to the elliptical formation with a 10 km separation experiencing moderate ionospheric effect. Figure 7.15 shows the relative position and velocity accuracies achieved. Comparing the results with shorter separations, the relative position errors are much more sensitive to the ionospheric change when the separation is extended to a medium length of 10 km. Overall the RW model is performing better in the position domain, but the HE model is still much more accurate for estimating the velocities. Analyzing the cross-track position errors closely, there is a period where the two solution do not concede, which occurs between the 2$^{\text{nd}}$ to 3$^{\text{rd}}$ hour time slot. This discrepancy is related to the ambiguity resolution state during the simulation, see Figure 7.16. The initial ambiguity resolution time for the

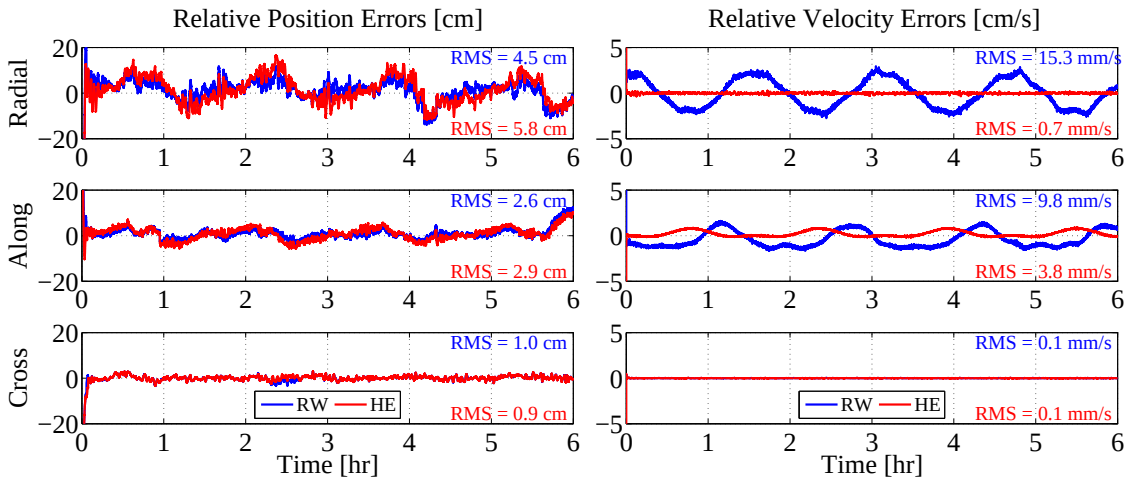


Figure 7.15: *Test 6 relative position and velocity errors.*

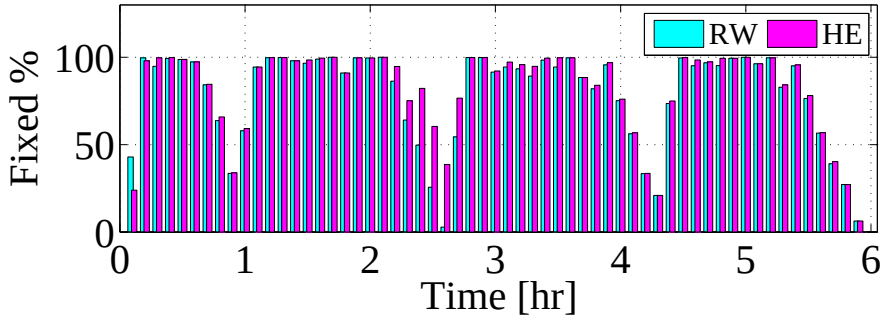


Figure 7.16: *Test 6 fixed ambiguity percentages.*

RW approach is 3 m19 s and 4 m34 s for the HE approach. The largest difference in the ambiguity fixed percentage between the RW and HE models is found during the 2nd to 3rd hour time slot. With the HE, the filter is able to maintain resolution for the new incoming ambiguities during the ionospheric peak period. However, the RW is unable to do so and integer ambiguity resolution actually failed. Another incident where failure of ambiguity resolution occurred is during the last ionospheric peak near the end of the simulation. During this peak period, both approaches failed to resolve any ambiguities and the fixed percentage dropped to zero. This is reflected

by the significant increase in the along-track errors during the last half an hour. In this situation, the HE seems to be performing better but more analysis is needed to verify the validity of this observation.

Summary

In this section, a performance analysis was conducted to quantify the relative navigation accuracies with the combinations of formation sizes and ionospheric conditions. The results for each of the test scenarios and propagation models are summarized Figure 7.17. The plots show the 3D RMS position and velocity errors for each of the

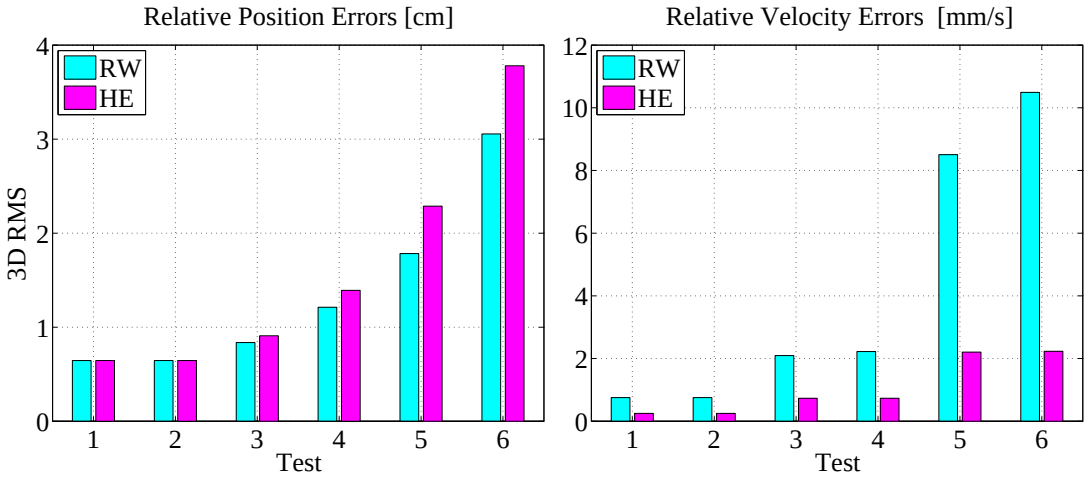


Figure 7.17: 3D RMS *errors of various test scenarios.*

test scenarios using both the RW and HE as process models. Based on the performance analysis, the following conclusions can be made from the investigation:

- Relative position accuracy was similar between the RW and HE approaches when the separation distances are relatively short. Even when the ionospheric errors were increased, the differential process was still able to eliminate its

effect. However, as the separation distance increased, the positioning performance of the HE was much more sensitive to the ionospheric condition. Overall the positioning performance obtained by both approaches was still reasonable considering that centimetre level accuracy was achieved for all the scenarios tested.

- By including the relative dynamics into the propagation model based on the HE, the relative velocity was significantly improved compared to the RW approach. Even when the ionospheric errors were increased, there was little impact on the HE approach. Only a minor degradation was observed for the test scenarios with medium separation (10 km). For cases where the separations were relatively short (0.5 km), even sub-millimetre per second accuracy was achieved. However, minor systematic errors were still visible from the relative velocity errors, which resulted from neglecting the J_2 perturbation.

7.2.2 Cold-Start Analysis

In the previous analysis, the fix time for integer ambiguity resolution was briefly analyzed. In some cases, large ambiguity resolution fix time differences between the two approaches were found. To address this problem, a more thorough investigation on the ambiguity resolution process from initial acquisition (cold-start) is conducted in this section. In cold-start analysis, the position and velocity is unknown. These information are estimated when signal acquisition begins.

To examine the ambiguity resolution process at different location of the 2-by-1 ellipse, the ambiguity resolution process is restarted every 10 minutes and re-initialized

with differential pseudorange measurements. A total of 36 ambiguity resolution trials were generated to reflect on the cold-start performance during different phases of the formation.

Results

The cold-start ambiguity resolution results for Test 1 to 4 can be found in the Appendix, Figure A.12 to Figure A.15. The results for Test Scenario 5 is shown in Figure 7.18. The plot shows the resolution of L_1 and WL ambiguity fix times for the

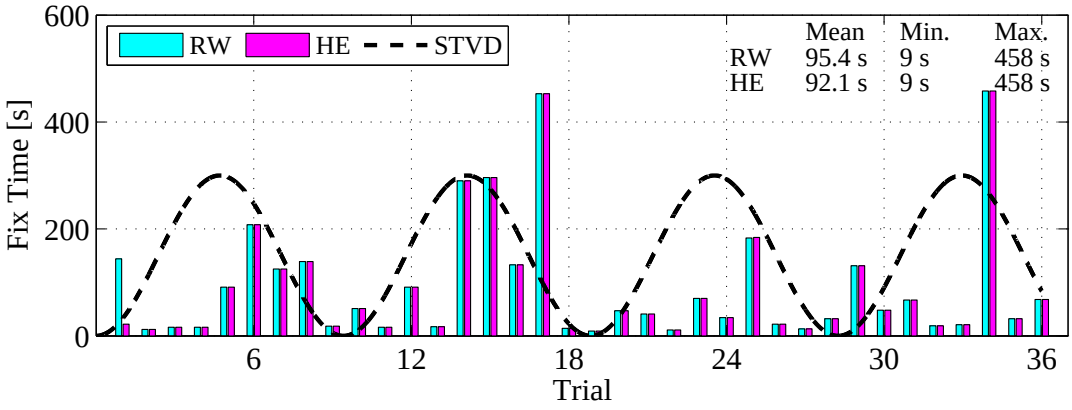


Figure 7.18: *Test 5 cold-start ambiguity resolution fix times.*

36 trials. The resolution performance of both process models are nearly the same. Only minor differences are found in the initial resolution trial and all other fix times are identical. The average resolution time is 95.4 s and the fix times varies from 9 s to 458 s. This large range is due to the variation of the ionospheric errors as it is the dominating factor effecting the GPS measurements. To illustrate this, the Scaled and Translated Vertical Displacement (STVD) is also included in the figure (black dashed line). In Section 5.3, the correlation between the vertical displacement and ionospheric errors have already been shown (Figure 5.6). Here, the result shows

that the cold-start ambiguity resolution times are somewhat correlated with the vertical displacement, which shows the impact of the ionospheric error on cold-start ambiguity resolution.

To examine the effect of an increased ionospheric error, the cold-start fix time for Test 6 is shown in Figure 7.19. Compared to the results from Test 5, a similar fix

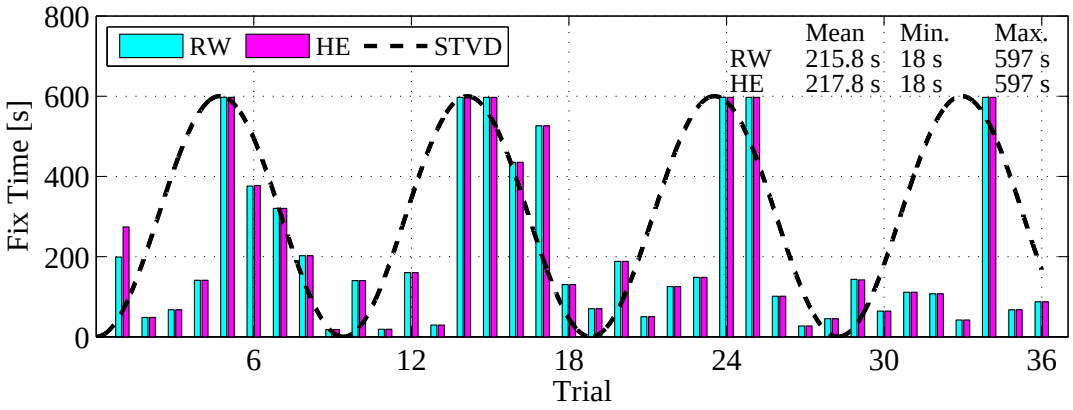


Figure 7.19: *Test 6 cold-start ambiguity resolution fix times.*

time pattern is found in Test 6. The major difference is that the fix times increased, especially during the ionospheric peak periods where fix times reached almost 10 minutes. Again, much correlation between the fix times and STVD is evidence. This indicates that the integer ambiguity resolution process is directly influenced by the level of ionospheric errors.

To analysis the integer ambiguity resolution pattern in terms of the relative satellite motion, the cold-start fix times are plotted on the 2-by-1 ellipse shown in Figure 7.20. . The figure shows that the integer ambiguity resolution times are much higher in the lower part of the ellipse, which is also where the peak ionospheric errors always occurs. This analysis precisely identifies where the cold-start ambiguity resolution

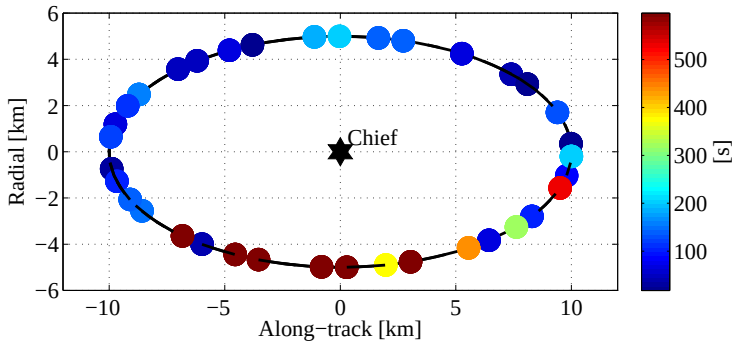


Figure 7.20: *Test 6 cold-start ambiguity resolution fix times at associated location.*

process is most impacted in terms of the satellite relative motion. It also provided a brief assessment on what the expected resolution time will be at different location on the ellipse.

A summary of the results obtained from the cold-start analysis for all test scenarios is shown in Table 7.6. The values shown are the mean, minimum and maximum

Table 7.6: *Statistics for cold-start ambiguity resolution analysis in seconds.*

Test	1		2		3		4		5		6	
Model	RW	HE	RW	HE	RW	HE	RW	HE	RW	HE	RW	HE
Mean	16.4	16.4	16.6	16.6	19.9	19.9	23.3	23.3	95.4	92.1	215.8	217.8
Min.	7	7	7	7	7	7	7	7	9	9	18	18
Max.	66	66	70	70	90	90	110	110	458	458	597	597

ambiguity fix times for the 36 trials for each of the test scenarios. This variation is proportional to the inter-satellite separation distances and ionospheric conditions. Depending on the situation, the minimum fix time can vary from 10 to 20 s and it is possible for the maximum to vary from 1 to 10 minutes. The cold-start ambiguity resolution performance is nearly the same from both the RW and HE approaches. Only minor differences are found when the separation are at medium lengths. This

analysis indicates that the level of process noise has little effect on the cold-start ambiguity resolution process. Instead, it is more dependent on the initial relative error level.

7.2.3 Signal Outage Analysis

In this section a signal outage analysis is conducted to investigate the relative navigation performance when no GPS measurements are available. This usually occurs when the signal-to-noise ratio becomes too low causing signal acquisition and tracking to fail. During this condition, the navigation solution will have to rely on the process model to propagate position and velocity over time. To investigate the performance of the RW and HE models during signal outages, Test Scenario 2, 4 and 6 are chosen for the analysis. These test scenarios were selected because the signal-to-noise ratio degraded by the moderate ionospheric conditions. Signal outages on all GPS satellites are induced every 10 minutes throughout the six hours of simulation. Different signal outage durations of 10, 20, 40, 60 and 90 s are used to analyze the robustness of the process model when no GPS information is available. In warm-start analysis, the relative position and velocity is known with certain accuracies. With these information available, integer ambiguity resolution is usually better than cold-start.

Results

The relative position and velocity errors during the 10 s signal outage are shown in Figure 7.21. The position errors plotted in black represents the epoch when no GPS measurements are available. By comparing the results between the RW and HE

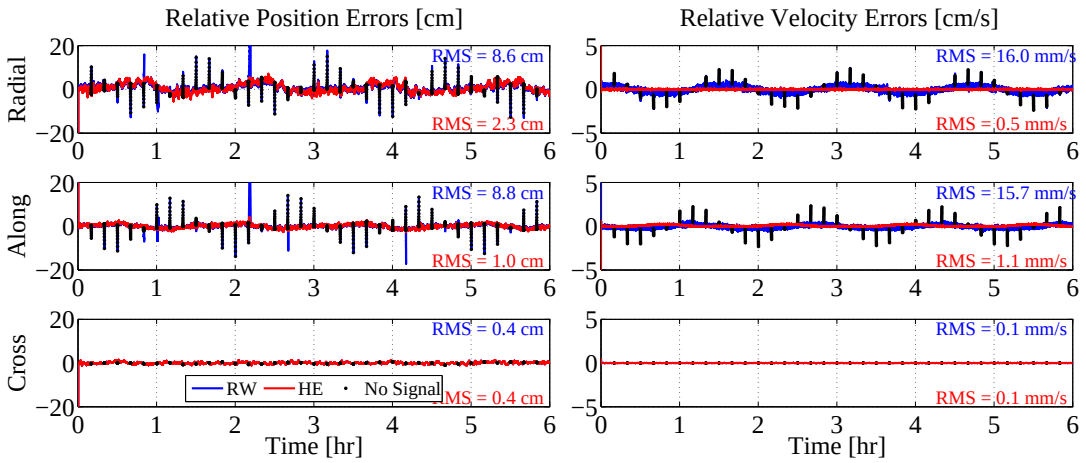


Figure 7.21: *Test 4 relative position and velocity errors with 10 s signal outages.*

approaches it is obvious that much better performance is demonstrated by the HE in both the position and velocity domains. The performance of the HE is nearly the same as the results obtained in Section 7.2.1 where no signal outage is induced. During a signal outage, the navigation solution relies solely on the process model, and the RW model is unable to estimate the relative satellite motion accurately. Depending where the outage is induced, the position errors in the radial and along-track directions may reach as much as 10 cm. Similar performance is found in the cross-track direction as both approaches can model zero motion.

To analyze the navigation performance for the different signal outage durations, the 3D RMS errors during outages are plotted in Figure 7.22. By including the relative satellite dynamics into the propagation model, the performance is clearly more robust than relying solely on the kinematics. For the RW approach, the 3D RMS already exceeds the centimetre level when the signal outage duration is 20s. The error level even reaches the metre level when the outages are 40 s. As for the HE model, even when the outage duration is extended to 90 s, the relative position

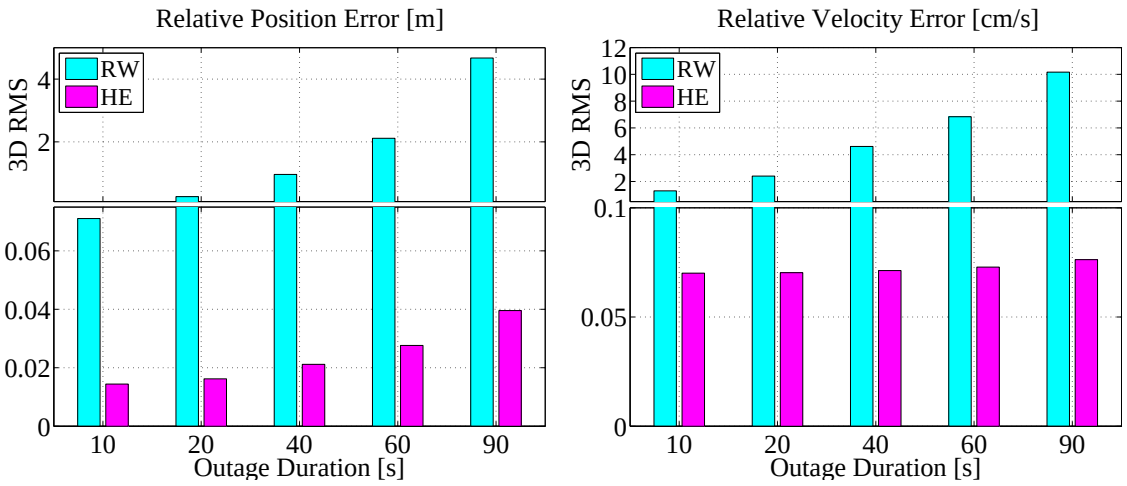


Figure 7.22: *Test 4* 3D RMS *errors during signal outages.*

error is always maintained at the millimetre level accuracy. In the velocity domain, the HE can maintain millimetre per second accuracies for all outage durations tested but the RW model was only able to provide centimetre per second level accuracy.

To analyze the performance in the ambiguity domain, the ambiguity fix times after signal outages of 10 and 20 s are plotted in Figures 7.23 and 7.24, respectively.

Referring to Figure 7.22, the 3D RMS position errors for 10 and 20 s outages are

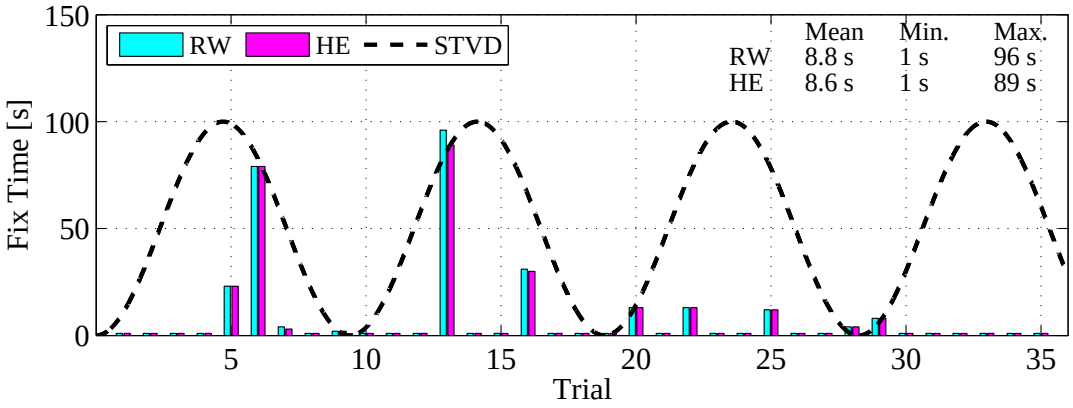


Figure 7.23: *Test 4 warm-start ambiguity fix times after 10 s signal outages.*

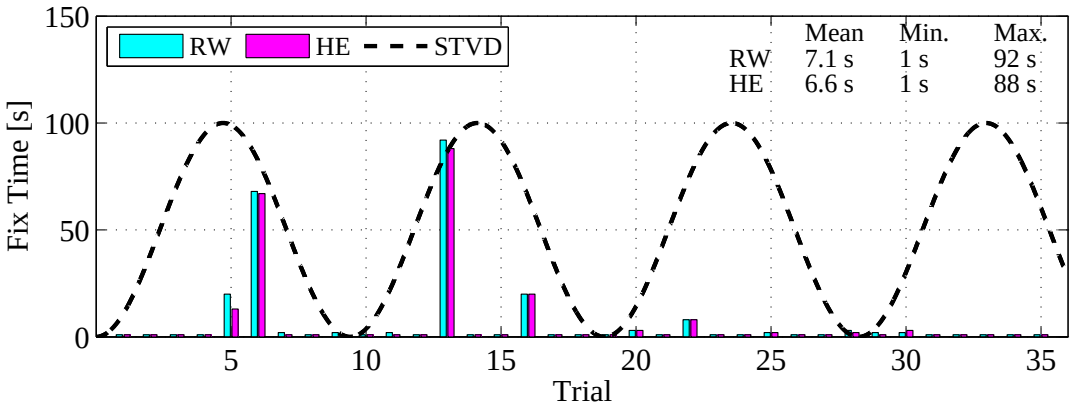


Figure 7.24: *Test 4 warm-start ambiguity fix times after 20 s signal outages.*

within the wavelength of WL. When the position errors are small, the correct integer solution set is obvious to the estimation process. Therefore, ambiguity resolution can be completed quickly and common performance can be observed between the two approaches. For most trials, ambiguity resolution is achieved within 5 s. There are only several trials where the fix time exceeds 30s and those trials were done during the signal outages from ionospheric peak periods.

Figures 7.25, 7.26 and 7.27 correspond to the warm-start ambiguity fix time for signal outage durations of 40, 60 and 90 s. For these outage durations, the 3D RMS errors of the RW approach exceeds the wavelength of WL. In this situation, better warm-start ambiguity resolution performance is demonstrated by the HE. In most of the trials, the HE approach is able the resolve the ambiguities much faster than the RW model. The average fix time of the HE approach is smaller in all outage durations tested. The level of improvement varies accordingly from each outage durations. Improvement is most evidenced when the 3D RMS position error of the RW model exceeds the wavelength of WL. For 90 s outages, the improvement achieved

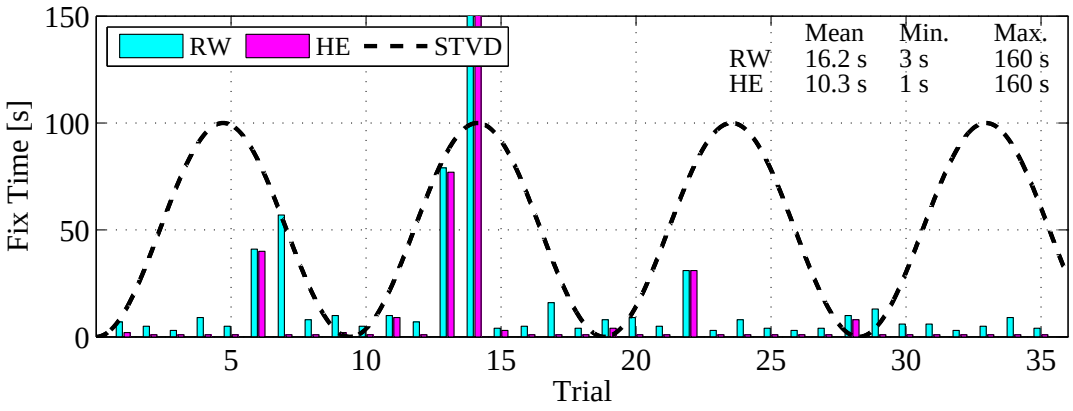


Figure 7.25: *Test 4 warm-start ambiguity fix times after 40 s signal outages.*

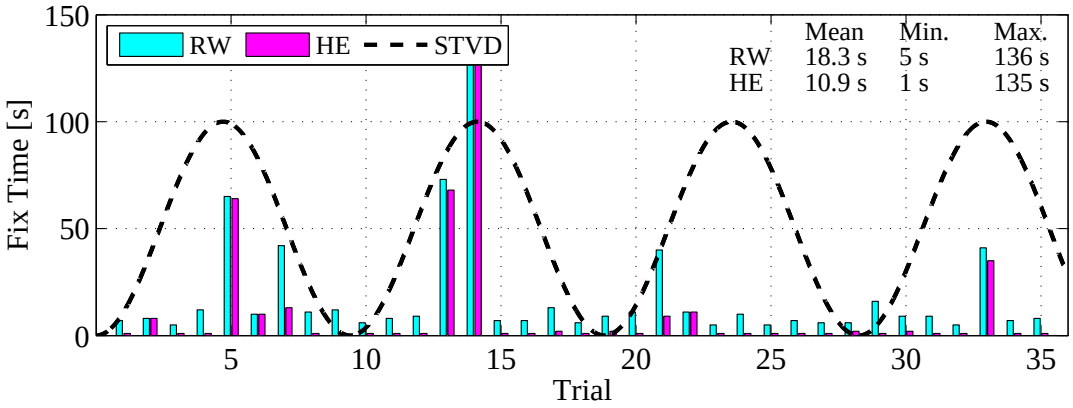


Figure 7.26: *Test 4 warm-start ambiguity fix times after 60 s signal outages.*

is nearly 35%.

A summary of the warm-start ambiguity fix times for Test Scenario 4 can be found in Table 7.7. The results show that by including the relative satellite dynamics into the process model the navigation robustness is strengthened during signal outage periods. Not only will the relative position and velocity accuracies be improved, ambiguity resolution after signal reacquisition will also be shortened. The results for Test Scenario 2 and 6 can be found in the Appendix, where similar findings are

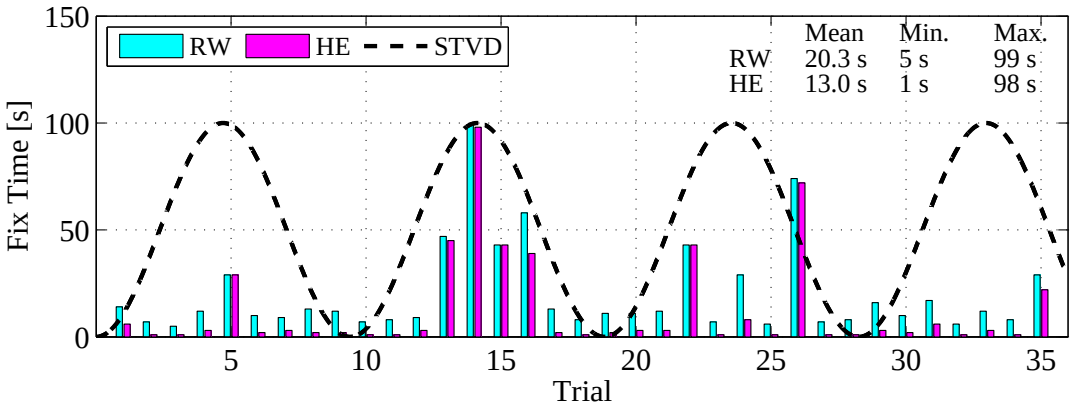


Figure 7.27: *Test 4 warm-start ambiguity fix times after 90 s signal outages.*

Table 7.7: *Test 4 warm-start ambiguity resolution fix time statistics in seconds.*

Outage	10 s		20 s		40 s		60 s		90 s	
Model	RW	HE	RW	HE	RW	HE	RW	HE	RW	HE
Mean	7.1	6.6	8.8	8.6	16.2	10.3	18.3	10.9	20.3	13.0
Min.	1	1	1	1	3	1	5	1	5	1
Max.	92	88	96	89	160	160	136	135	99	98

observed.

Chapter 8

Conclusions & Recommendations

8.1 Conclusions

This research focused on several practical and important aspects of satellite formation flying. The thesis began with a thorough investigation of relative satellite motion using the HE. Based-on the HE, two configuration types (2-by-1 elliptical and circular formation) were studied to evaluate their possibility for future geodetic formation flying missions. The results of a theoretical analysis indicated that the 2-by-1 elliptical formation is actually immune to the J_2 perturbation, which affirms that the formation will not drift apart over time due to such a perturbation. The important benefit of this finding is that less fuel is required for formation control or adjustment, which consequently extends the life span of the mission allowing for more valuable information to be collected and ultimately reduces the overall cost. Another formation type investigated was the circular formation. It was found that the circular configuration is not predominately J_2 invariant. Only when certain constraints are applied will the formation become J_2 invariant, however, these conditions will create impracticalities in geodetic missions. Furthermore, the accuracy of the HE for modelling satellite motion was thoroughly assessed. This analysis was designed and executed with the absolute and relative motions in both the central and J_2 disturbed gravitational fields under consideration. The analysis indicated that the

HE worked extremely well in the central gravitational field, however, the accuracy degraded rapidly over time in the J_2 disturbed gravitational field, which suggests that the HE is only valid for short durations in realistic applications.

The thesis discussed GPS conditions in a LEO spaceborne environment. The simulation software, SIMGNSSII^{TM}, was presented and the algorithms used to replicate the GPS observations and errors were thoroughly discussed. Following the discussion was the relative navigation algorithms for precise relative orbit determination. Two propagation approaches based-on the relative satellite kinematics (Random Walk velocity model) and dynamics (HE) were introduced. The performances of the two approaches were carefully analyzed using different test scenarios. The initial assessment began with a sensitivity analysis, where the spectral densities of the HE were found to be two magnitudes smaller than the Random Walk velocity model (RW). This was an indication that better performance should be expected from the HE. The investigation continued with a performance comparison, where the RW and HE models were combined with different measurement combinations to determine the optimal combination for further analysis. Through the performance comparison, it was determined that the optimal combination was using the L_1 and Widelane (WL) carrier, therefore strategies containing these measurements were selected for the performance test as well as the cold-start and warm-start ambiguity analyses.

There were several limitations to the research conducted. The first was in the simulation of satellite orbit, where higher order gravitation, atmospheric drag and solar pressure were neglected. Although these disturbances are not as forceful as J_2, but their cumulative affect over long periods of time can be significant. Furthermore, the analysis conducted was limited to equatorial orbits, which minimized the effect of

J_2 on the satellite relative motion. Another limitation was due to the linearity of the HE, so only short and medium formations were analyzed. Extending the separation any longer would have caused the reduced dynamic model to fail.

Based-on the test results achieved from this chapter, the following conclusions can be drawn:

- The RW and HE approaches demonstrated their ability to provide precise position accuracy at the centimetre level when carrier phase measurements are used with proper ambiguity resolution techniques. The results achieved are similar to the work of Busse (2003) and satisfied the navigation requirements of most geodetic missions. For formations with relatively short separation distances (0.5 km), a similar positioning performance was found between the two approaches. However, as the separation was increased to a medium level, it appears that the RW was more suitable as it demonstrated slightly better accuracy. The poorer performance of the HE was related to the increase of approximation error due to the linearization of the relative satellite dynamics. Nevertheless, the consideration of the dynamics did significantly improve the velocity estimation, which is beneficial for formation flying control (How & Tillerson, 2001).

- The cold-start ambiguity resolution performance was shown to be affected by the relative GPS error level between the formation flying spacecraft, which are mainly dependent upon the instantaneous ionospheric conditions and inter-satellite distances. Although, the HE model was able to reduce the spectral densities by two orders of magnitudes, the analysis showed little ambiguity

resolution difference between the two approaches. This was because the initial positioning errors were too significant, where the reduction of process noise had hardly any effect on ambiguity resolution.

- During a complete signal outage, the RW approach was shown to provide metre level accuracy (3D RMS) for outages lasting up to 90 s. For the same outage duration, the accuracy of the HE was maintained at the centimetre level, and thus presented a significant advantage over the RW approach. As for the accuracy of the velocity estimates, the HE showed little sensitivity to the outage during and the 3D RMS errors were kept in the sub-millmetre per second level. This was, however, not the case for the RW approach, as the velocity accuracy varied in proportion to the outage duration. For outages lasting 90 s, the 3D RMS error can exceed the decimetre per second level.

- The warm-start integer ambiguity performance was shown to be directly related to the position accuracy of the process model during signal outage. During outages, if the 3D RMS error level was kept under the wavelength of the WL, then the warm-start ambiguity fix times were similar between the RW and HE approaches. However, when the outages were increased to the point where the error of the RW model exceeded 86 cm, the HE demonstrated significant ambiguity resolution improvements, since it was able to keep the position error level well under the WL wavelength during outages.

- Comparing the navigation performance achieved by GPS to other sensors, such as the K-band laser ranging system, the accuracy is certainly not comparable, since laser ranging can provide precise range measurements at micrometre level

(Rodrigues et al., 2003). Though keep in mind, for formations limited to two spacecrafts, laser ranging can only provide measurements in 1-dimension. This has been demonstrated by the GRACE mission where the along-track motion has been precisely monitored by K-band laser. Only when three or more spacecrafts are incorporated, then inner constraints can be applied to resolve for their relative positions. However, another sensor is still required for orienting the configuration.

8.2 Recommendations

Based on the results and the conclusions, the following recommendations can be made to address the limitations of this research:

- For the Kalman filter implemented, dual-frequency measurements were used for relative orbit determination. However, the information within these measurements were not fully exploited to estimate the relative ionospheric effect, which was shown to be the dominating error source for medium separation formations. The ionosphere parameter should not be grouped together with other residuals and instead it should be modelled as additional states within the Kalman filter. Such an implementation should improve the relative navigation performance significantly, especially during ionospheric peak periods. Another alternative is using the Ionopheric Free (IF) observables. Such a method can reduced the ionospheric error, however noisier navigation performance will be expected.

- Although much effort has been done to realistically replicate the spaceborne GPS environment. There are still limitations to a software simulator. In order to model the GPS errors more accurately, a suggestion is to use the Spirent hardware simulator, where it can also be used to assess the performance of a GPS receiver which has been demonstrated by Montenbruck (2003) to analysis the NovAtel OEM4-G2L receiver.

- The multipath model used in this research was reasonable for spaceborne applications since the only available reflective source is the spacecraft itself. For the simulation conducted, local multipath errors were assumed to be uncorrelated between spacecraft. However, Kroes (2006) illustrated that multipath errors of the GRACE spacecraft are somewhat correlated, but no values were provided to quantify this relationship. More work is needed to address this issue of the multipath errors on a formation flying mission. It is speculated that there is a relationship with the relative attitude of the spacecraft as this is a key parameter that would effect the multipath geometry.

- For the simulations conducted, an important assumption made was the continuity and availability of the communication link between the spacecraft that allowed the transmission of the chief information to the deputy. In reality, this link or information may not be available, hence pre-planned procedures are required to handle such incidents. These procedures should guarantee a certain degree of accuracy during the absence of chief information. A real formation flying experiment, such as the Canadian Advanced Nanospace eXperiment (CAN-X), can be utilized to be verified the performance of these autonomous

procedures.

- To replicate the satellite motion more realistically, perturbations such as higher order gravitation, atmospheric drag and solar pressure should be included in the orbit simulation. These perturbations should also be considered in configuration design, since their cumulative affect over time can be significant. Furthermore, high inclination orbits should be studied to fully analyze the effect of J_2 on satellite relative motion. Such an analysis is important as high inclination orbits are preferred by geodetic missions for maximizing the spatial coverage.

- More accurate dynamic models should be studied and used for formations with long separations. A numerical integration approach was demonstrated by Kroes (2006). A complex dynamic model was incorporated into the Kalman filter to navigate the GRACE spacecraft relatively and sub-millimetre positioning accuracy was achieved.

References

Abidin H (1993) *Computational and Geometrical Aspects of On-The-Fly Ambiguity Resolution*. Ph.D. thesis, Department of Surveying Engineering, University of New Brunswick.

Adams J, Robertson A, Zimmerman K, How JP (1996) Technologies for Spacecraft Formation Flying. In: *Proceedings of the ION GPS Conference*.

Alfriend KT, Schaub H (2000) Hybrid Cartesian and Orbit Elements Feedback Law for Formation Flying Spacecraft. In: *Proceedings of the AIAA Astrodynamics Specialist Conference, Denver, Co*.

Alfriend KT, Schaub H, Gim DW (2000b) Gravitational Perturbations, Nonlinearity and Circular Orbit Assumption Effects on Formation Flying Control Strategies. In: *Proceedings of the AAS Guidance and Control Conference*.

Axelrad P, Enge P (1996) GPS Navigation Algorithms. In: B Parkinson, J Spiker (eds.), *Global Positing System: Theory and Applications*, vol. 1, pp. 409–433, American Institue of Aeronautics and Astronautics.

Bate RR, Mueller DD, White JE (1971) *Fundamentals of Astrodynamics*. Dover Publications.

Bauer FH, Hartman K, How JP, et al. (1999) Enabling Spacecraft Formation Flying through Spaceborne GPS and Enhanced Automation Technologies. In: *Proceedings of the ION GPS Conference, Nashville, TN*.

Braasch M, van Graas F (1991) Guidance Accuracy Considerations for Real time GPS Interferometry. In: *Proceedings of ION GPS-91*, pp. 373–386.

Brown RG, Hwang PYC (1997) *Introduction to Random Signals and Applied Kalman Filtering*. 3rd edn., John Wiley & Sons.

Bugoslavskaya N (1962) *Solar Activity and the Ionosphere, for Radio Communications Specialists*. Pergamon Press Ltd., New York.

Busse FD (2003) *Precie Formation-States Estimation In Low Earth Orbit Using Carrier Differntial GPS*. Ph.D. thesis, The Department of Aeronautics and Astronautics, Standford University.

Busse FD, How JP, Simpson J (2002b) Demonstration of Adaptive Extended Kalman Filter for Low Earth Orbit Formation Estimation Using CDGPS. In: *Proceedings of the ION GPS Conference, Portland, OR*.

Cannon M (1991) *Airborne GPS/INS with an Applications to Aerotriangulations*. Ph.D. thesis, Department of Survey Engineering, University of Calgary.

Cannon M, Lachapelle G (2004) *SimGNSSII*^{tm} *Operating Manual*. Department of Geomatics Engineering, University of Calgary.

Cartan Élie Joseph (1894) *Sur la structure des groupes de transformations finis et continus*. Ph.D. thesis, Paris, Nony.

Chen D, Lachapelle G (1994) A Comparison of the FASF and Least-Squares Search Algorithms for On-the-Fly Ambiguity Resolution. *Journal of the Institute of Navigation* **42**(2):371–390.

Clohessy WH, Wiltshire RS (1960) Terminal Guidance System for Satellite Rendezvous. *Journal of the Aerospace Sciences* **27**(9):653–658.

Corazzini T, Robertson A, Adams JC (1997) GPS Sensing for Spacecraft Formation Flying. In: *Proceedings of the ION GPS Conference, Kansas City, Missouri*.

Counselman C (1981) Miniature Interferometer Terminals for Earth Surveying: Ambiguity and Multipath with the Global Positioning System. *IEEE Transactions on Geoscience and Remote Sensing* **19**(4):244–252.

Crawford S (2005) *Performance Evaluation of Sensor Combinations for Mobile Platoon Control*. Master's thesis, Department of Geomatics Engineering, University of Calgary.

Dong L (2004) *SimGNSSII^{tm} Algorithm Design Document*. Tech. rep.

Ebinuma T (2001) *Precision spacecraft rendezvous using GPS: an integraded hardware approach*. Ph.D. thesis, University of Texus.

Feess WA, Stephens SG (1987) Evaluation of GPS Ionospheric Model. *IEEE Transactions on Aerospace and Electronic Systems* **AES-23**(3):332–338.

Foster J (2000) Quantitative investigation of ionospheric density gradients at mid latitudes. In: *Proceedings of ION National Technical Meeting*.

Frei E, Beutler G (1990) Rapid static positioning based on the Fast Ambiguity Resolution Approach "FARA": Theory and First Results. *Manuscripta Geodaetica* **15**:325–356.

Gao Y (2004) ENGO 629 Lecture Notes: Advance Estimation Methods and Analysis. Department of Geomatics Engineering, University of Calgary.

Gelb A, Kasper Jr JF, Nash Jr RA, et al. (1974) *Applied Optimal Estimation*. The M.I.T. Press, Cambridge, MA.

Georgiadou Y, Kleusberg A (1988) On Carrier Signal Multipath Effects in Relative GPS Positioning. *Manuscripta Geodaetica* **13**(3):172–179.

Gim DW, Alfriend KT (2001) The State Transition Matrix of Relative Motion for the Perturbed Non-Circular Reference Orbit. In: *Proceedings of the AAS/AIAA Space Flight Mechanics Meeting, Santa Barbara, CA,*.

Hatch R (1991) Ambiguity Resolution While Moving-Experimental Results. In: *Procceedings of ION GPS 91*, pp. 707–713.

Hill G (1878) Researches in the Lunar Theory. *American Journal of Mathematics* **1**:5–26.

Hopfield H (1970) *Tropospheric Effect on Electromagnetically Measured Range: Prediction from Surface Weather Data*. Tech. rep., Applied Physics Laboratory, Johns Hopkins, University, Baltimore, MD.

How J, Tillerson M (2001) Analysis of the Impact of Sensor Noise on Formation Flyin Control. In: *American Control Conference*.

ICD-GPS-200c (1997) Navstar GPS Space Segment / Navigation User Interfaces. U.S. Air Force.

IGS (2005) International GNSS Service Data and Products. URL `http://igscb.jpl.nasa.gov/components/compindex.html`

Julien O (2005) *Design of Galileo L1F Receiver Tracking Loops*. Ph.D. thesis, Department of Geomatics Engineering, University of Calgary.

Kaplan D (1996) *Understanding GPS Principles and Applications*. Artech House Publishers.

Kaula WM (2000) *Theory of Satellite Geodesy: Applications of Satellites to Geodesy*. Dover Publicatoins.

Klobuchar J (1996) Ionospheric Effects on GPS. In: B Parkinson, J Spilker (eds.), *Global Positioning System, Theory and Application*, vol. 1, pp. 485–514, American Institute of Aeronautics and Astronautics.

Klobuchar JA (1986) Design and Characteristics of the GPS Ionospheric Time-Delay Algorithm for Single-Frequency Users. In: *Proceedings of PLANS'86 - Position Location and Navigations Symposium*, pp. 280–286.

Kroes R (2006) *Precise Relative Positioning of Formation Flying Spacecraft using GPS*. Netherlands Geodetic Commision.

Leitner J, Bauer F, Folta D, et al. (2002) Formation Flight in Space. *GPS World* pp. 22–31.

Leung S, Montenbruck O (2005) Real-Time Navigation of Formation Flying Spacecraft Using GPS Measurements. *Journal of Guidance, Control, and Dynamics* **28**(2).

Liao X (2000) *Carrier Phase Based Ionosphere Recovery Over A Regional Area GPS Network*. Master's thesis, Department of Geomatics Engineering, University of Calgary.

Liu J (2003) *Implementation and Analysis of GPS Ambiguity Resolution Strategies in Single and Multiple Reference Station Scenarios*. Master's thesis, Department of Geomatics Engineering, University of Calgary.

Liu Z (2004) *Ionosphere Tomographic Modeling and Applications Using Global Positioning System (GPS) Measurements*. Ph.D. thesis, Department of Geomatics Engineering, University of Calgary.

Lu G (1995) *Development of a GPS Multi-Antenna System for Attitude Determination*. Ph.D. thesis, Department of Geomatics Engineering, University of Calgary.

Luo N (2001) *Precise Relative Positioning of Multiple Moving Platforms Using GPS Carrier Phase Observables*. Ph.D. thesis, Deparment of Geomatics Engineering, University of Calgary.

Massonnet D (2001) Capabilities and Limitations of the Interferometric CartWheel. *IEEE Trans Geosci Remote Sensing* **39**(3):506–520.

Misra P, Burke B, Pratt M (1999) GPS Performance in Navigation. In: *Proceedings of the IEEE*, vol. 87 of *1*.

Misra P, Enge P (2001) *Global Positiong System: Signals, Measurements and Performance*. Ganga-Jamuna Press, P.O. Box 692, Lincoln, Massachusetts 01773.

Montenbruck O (2003) *Performance Assessment of the NovAtel OEM4-G2 for LEO Satellite Tracking*. Space Flight Technology, German Space Operation Center (GSOC), Deutsches Zentrum fü Luft- und Raumfahrt (DLR).

Montenbruck O, Gill E (2000) *Satellite Orbits - Models, Methods and Applications*. Springer-Verlag, Berlin-Heidelberg.

Montenbruck O, Kroes R (2003) In-Flight Performance Analysis of the CHAMP BlackJack GPS Receiver. In: *EGS - AGU - EUG Joint Assembly*.

Moon Y (2004) *Evaluation of 2-Dimensional Ionosphere Models for National and Regional GPS Networks in Canada*. Master's thesis, Department of Geomatics Engineering, University of Calgary.

Moritz H (1980) *Advanced Physical Geodesy*. Wichmannn.

Newton I (1687) *Philosophiae Naturalis Principia Mathematica*. Royal Society of London.

Olsen EA, Park CW, How JP (1998) 3D Formation Flight Using Differential Carrier-Phase GPS Sensors. In: *Proceedings of the ION GPS Conference, Nashville, TN*.

Olynik M (2002) *Temporal Characteristics of GPS Error Sources and Their Impact on Relative Positioning*. Master's thesis, Deparment of Geomatics Engeering, University of Calgary.

Park C, How J, Capots L (2000) Sensing Technologies for Formation Flying Spacecraft in LEO Using CDGPS and Inter-Spacecraft Communications System. In: *Institute of Navigation GPS-00 Conference*, pp. 1595–1607.

Parkinson B (1996) GPS Error Analysis. In: B Parkinson, J Spilker (eds.), *Global Positioning System: Theory and Applications*, vol. 1, Global Positioning System: Theory and Applications.

Petovello M, Cannon M, Lachapelle G (2000) *Combined Code and Carrier for NAVigation with GPS and GLONASS (C^3NAVG^2) Operation Manual*. Department of Geomatics Engineering, The University of Calgary, 2500 University Drive NW, Calgary, AB, Canada.

Ray J (2000) *Mitigation of GPS Code and Carrier Phase Multipath Effects Using a Multi-Antenna System*. Ph.D. thesis, Department of Geomatics Engineering, University of Calgary.

Richert T (2003) *Kinematic Positioning Inside the Ionosphere*. Tech. rep., Waypoint Consulting Inc.

Rishbeth H, Garriott O (1969) *Introduction to Ionospheric Physics*. Academic Press Inc., New York.

Rodrigues M, Foulon B, Liorzou F, Touboul P (2003) Flight Experience on CHAMP and GRACE with Ultra-sensitive Accelerometers and Return for LISA. In: *Classical and Quantum Gravity*, vol. 20 of *10*, pp. S291–S300(1), Institute of Physics Publishing.

RTCA (1998) Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment. RTCA/DO-229A, SC-159.

Rummel R (1986) Satellite Gradiometry. In: H Sünkel (ed.), *Mathematical and Numerical Techniques in Physical Geodesy*, vol. 7, pp. 317–364, Springer-Verlag Berlin Heidelberg.

Schaub H (2002) Spacecraft Relative Oribt Geometry Description Through Orbit Elements Differeneces. In: *14th U.S. National Congress of Theoretical and Applied Mechanics*.

Schaub H, Alfriend KT (1999) J_2 Invariant Relative Orbits for Spacecraft Formations. *Journal of Celestial Mechanics for Publication* .

Schaub H, Junkins JL (2003) *Analytical Mechanics of Space Systems*. AIAA Education Series, American Institute of Aeronautics and Astronautics, Inc., 1801 Alexander Bell Drive, Reston, VA 20191-4344.

Schwarz K (1999) Fundamentals of Geodesy. Department of Geomatics Engineering, University of Calgary.

Schweighart SA, Sedwick RJ (2001) A Perturbative Analysis of Geopotential Disturbances for Satellite Cluster Formation Flying. In: *Proceedings of the IEEE Aerospace Conference*.

Schweighart SA, Sedwick RJ (2002) A High Fidelity Linearized J_2 Model for Satellite Formation Flight. *AIAA Journal* .

Seeber G (2003) *Satellite Geodesy*. 2nd edn., Walter de Gruyeter GmbH & Co. KG, 10785 Berlin.

Shrestha SM (2003) *Investigations into the Estimation of Tropospheric Delay and Wet Refractivity Using GPS Measurements*. Master's thesis, Department of Geomatics Engineering, University of Calgary.

Skone S (1998) *Wide Area Ionosphere Grid Modelling in the Auroral Region*. Ph.D. thesis, Department of Geomatics Engineering, University of Calgary, Calgary, AB.

Sneeuw N (2000) *A Semi-Analytical Approach to Gravity Field Analysis from Satellite Observations*. Ph.D. thesis, Deutsche Geodätische Kommission.

Sneeuw N (2002a) ENGO 427 Lecture Notes: Gravity Theory (Physical Geodesy). Department of Geomatics Engineering, University of Calgary.

Sneeuw N (2002b) LISA/Cartwheel Orbit Type for Future Gravity Field Satellite Missions. In: *Proceedings of the Weikko A. Heiskanen Symposium in Geodesy*.

Spilker Jr J (1996) *Global Positioning System: Theory and Applications*, vol. 1. Stanford Telecom, Sunnyvale, California.

Sünkel H (1998) Dynamic Satellite Geodesy: Theory and Application. Institude of Theoretical Geodesy, Department of Mathematical Geodesy and Geoinformatics, Technical University Graz.

Teunissen P (1995) The Least squares Ambiguity Decorrelation Adjustment: A method for Fast GPS Integer Ambiguity Estimation. *Journal of Geodesy* **70**:65–82.

Teunissen P (1998) Success Probability of Integer GPS Ambiguity Round and Bootstrapping. In: *Journal of Geodesy*, vol. 72, pp. 606–612, Springer-Verlag.

CODE (2006) Global Ionosphere Maps Produced by Centre for Orbit Determination in Europe.
URL http://www.aiub.unibe.ch/ionosphere.html

USCG (2002) U.S. Coast Guard Navigation Center GPS Almanac Information.
URL http://www.navcen.uscg.gov/gps/almanacs.htm

Vadali S, Young X, Schaub H, Alfriend KT (1999) Fuel Optimal Control for Formation Flying of Satellites. AIAA Guidance, Navigation and Control Conference.

Vadali SR, Vaddi SS, Alfriend KT (2000) A Intelligent Control Concept for Formation Flying Satellite Constellations. *International Journal of Robust and Nonlinear Control* .

Van Dierendonck A, Fenton P, Ford T (1992) Theory and Performance of Narrow Correlator Spacing in a GPS Receiver. *Navigation* **39**(3):265–283.

Wanninger L (1993) Effects of the Equatorial Ionosphere on GPS. *GPS World* .

Ward P (1996) Satellite Signal Acquisition and Tracking. In: E Kaplan (ed.), *Understanding GPS Principles and Applications*, pp. 119–208, Artech House, Boston.

Welch P (1967) The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging over Short, Modified Periodograms. *IEEE Transactions on Audio Electroacoustics* **AU-15**:70–73.

Wells D, Beck N, Delikaraoglou D, et al. (1986) Guide to GPS Positioning. Canadian GPS Associates, Fredericton.

Wu H (2003) *On-The-Fly GPS Ambiguity Resolution with Inertial Aiding*. Master's thesis, Department of Geomatics Engineering, University of Calgary.

Xu C, Tsoi R, Sneeuw N (2004) Satellite Formation Flying for Geodetic Applications: Analysis of J_2 Disturbed Orbits. In: *Proceedings of the 2004 Joint Assembly, Montreal, QE.*

Xu C, Tsoi R, Sneeuw N (2005) Analysis of J2-Perturbed Relative Orbits for Satellite Formation Flying. In: C Jekeli, L Bastos, J Fernandes (eds.), *Gravity, Geoid and Space Missions*, vol. 129 of *International Association of Geodesy Symposia*, Springer Berlin Heidelberg.

Zimmerman KR, Cannon Jr RH (1995) Experimental Demonstration of GPS for Rendezvous Between Two Prototype Space Vehicles. In: *Proceedings of the ION GPS Conference, Palm Springs, CA.*

Appendix A

Appendix

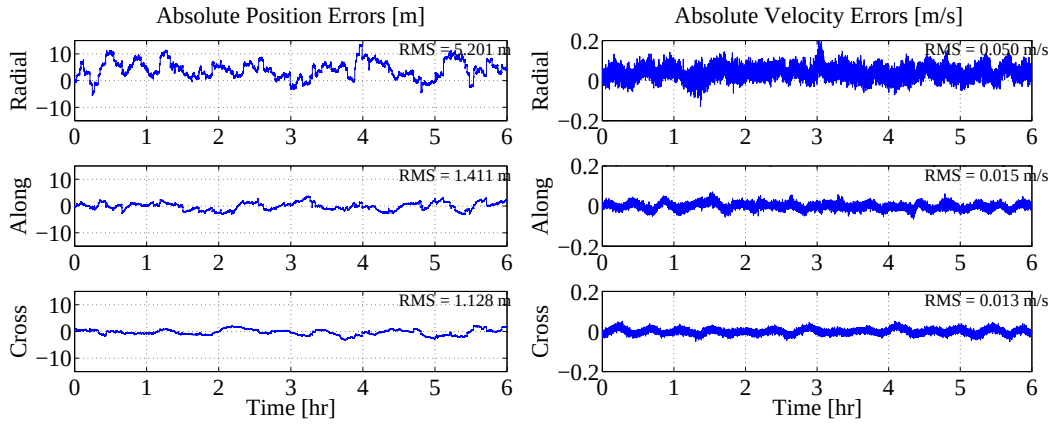
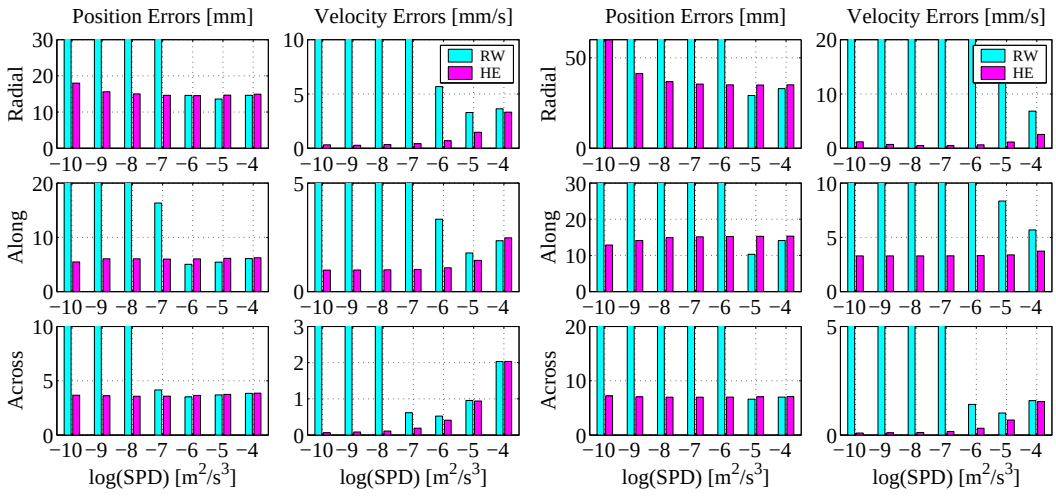


Figure A.1: *Absolute position and velocity errors derived from* C^{3}NAVG$^{2\text{TM}}$.



(a) 3 km separation (b) 10 km separtion

Figure A.2: *Sensitivity analysis for elliptical formations*

Performance Comparison

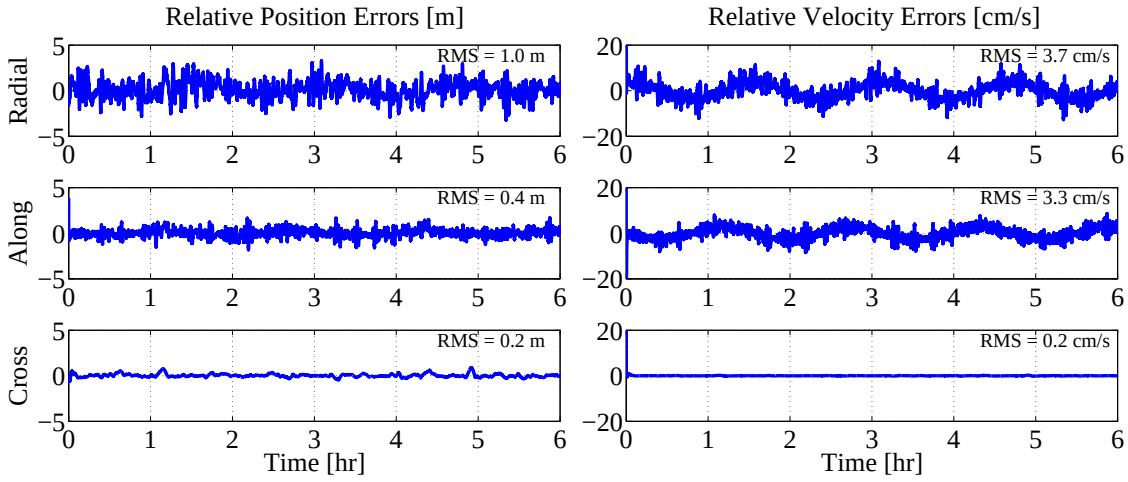


Figure A.3: *Strategy 1 relative position and velocity errors.*

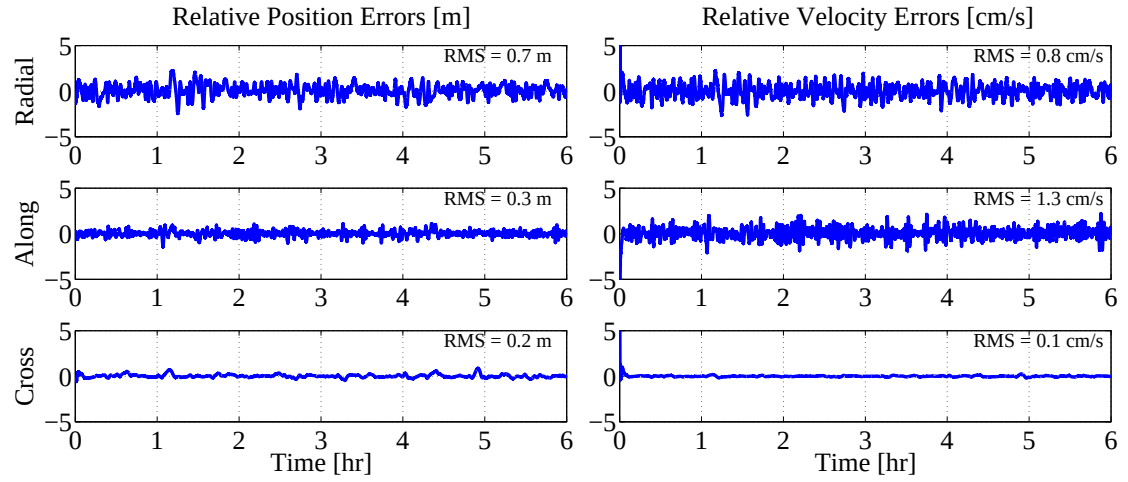


Figure A.4: *Strategy 2 relative position and velocity errors.*

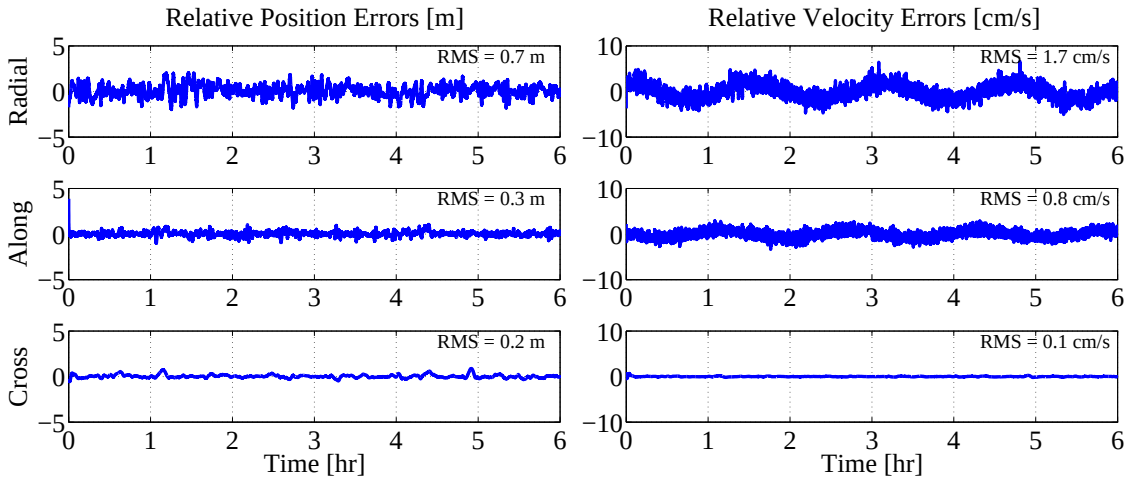


Figure A.5: *Strategy 3 relative position and velocity errors.*

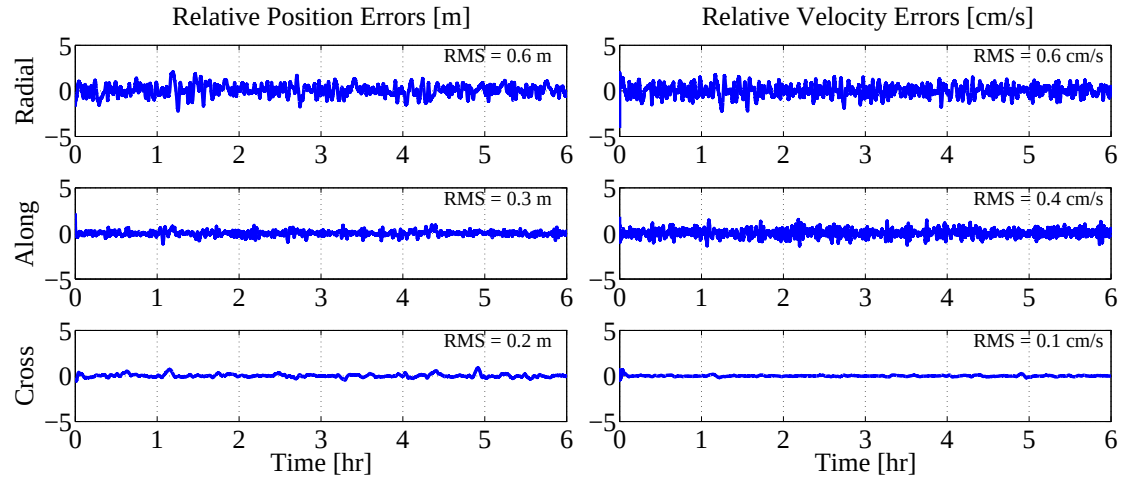


Figure A.6: *Strategy 4 relative position and velocity errors.*

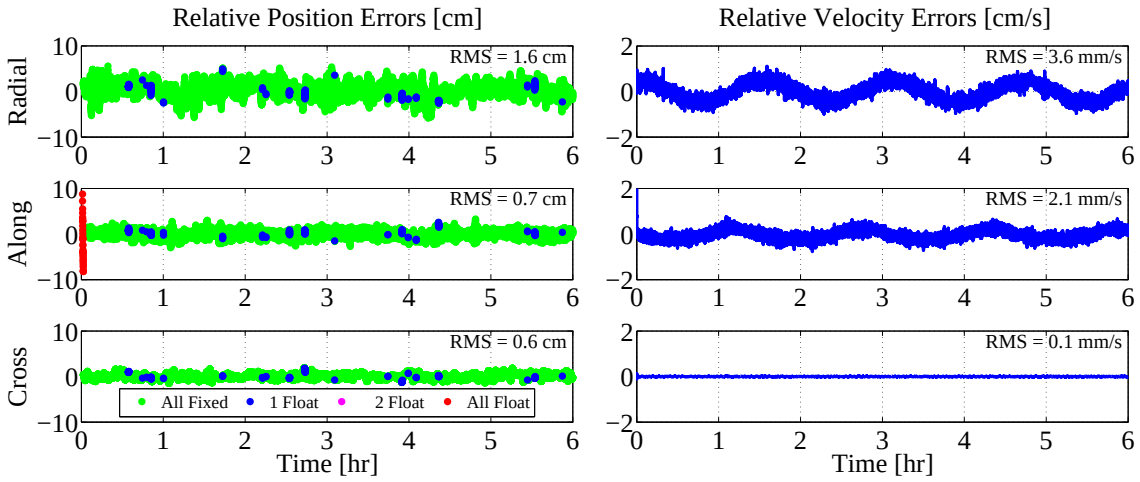


Figure A.7: *Strategy 5 relative position and velocity errors.*

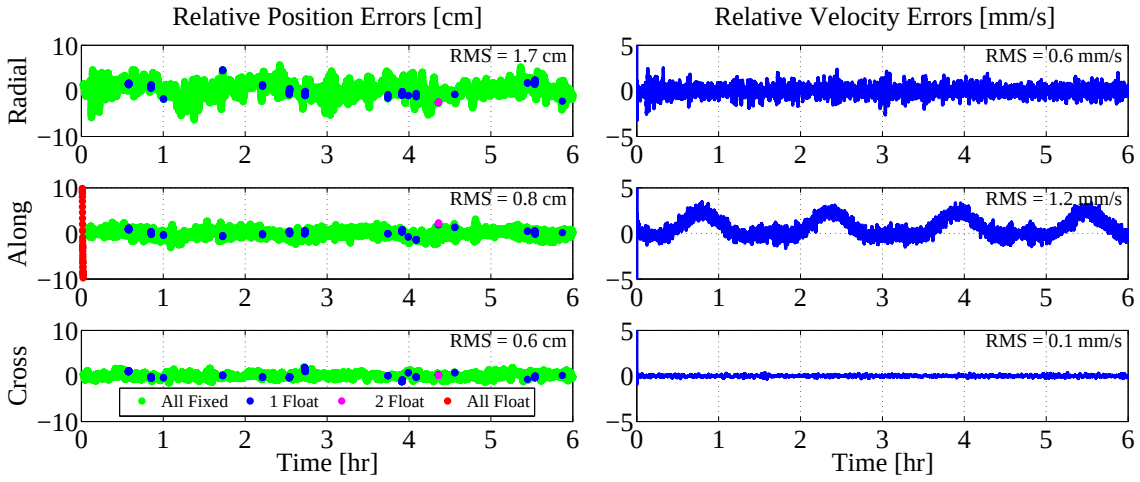


Figure A.8: *Strategy 6 relative position and velocity errors.*

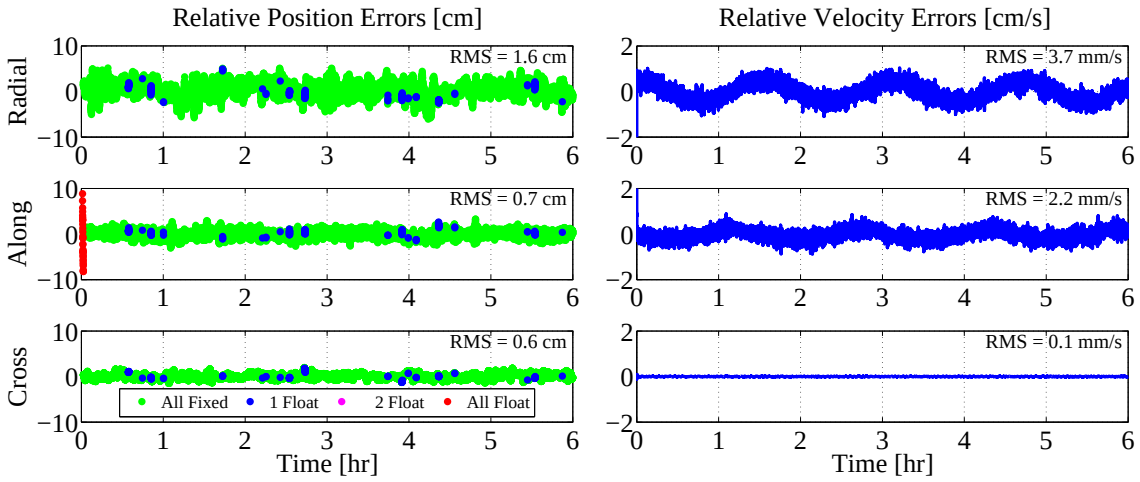


Figure A.9: *Strategy 7 relative position and velocity errors.*

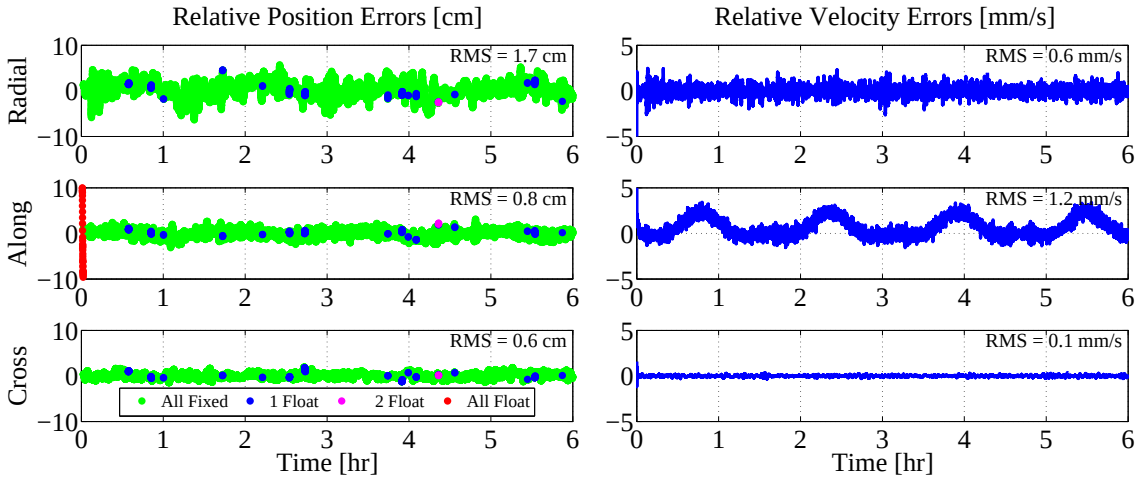


Figure A.10: *Strategy 8 relative position and velocity errors.*

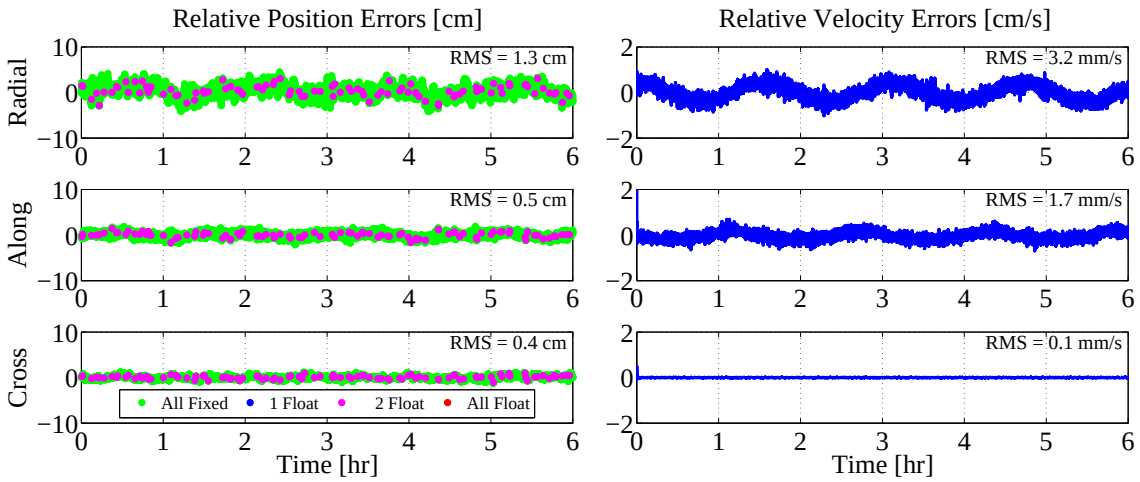


Figure A.11: *Strategy 9 relative position and velocity errors.*

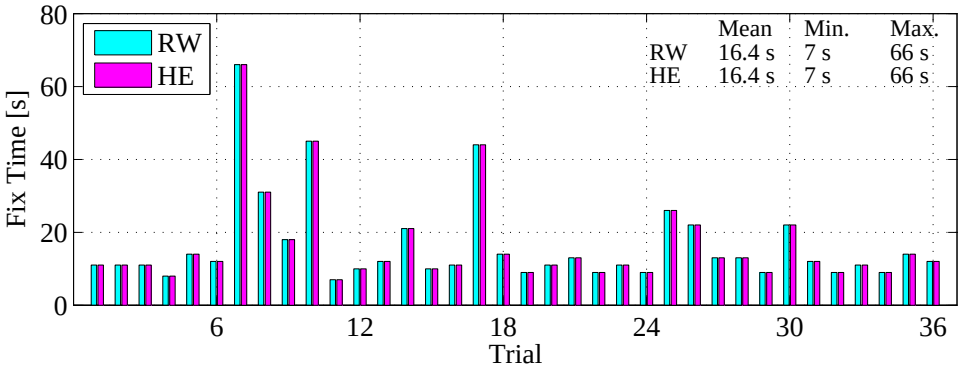


Figure A.12: *Test 1 cold-start ambiguity resolution fix times.*

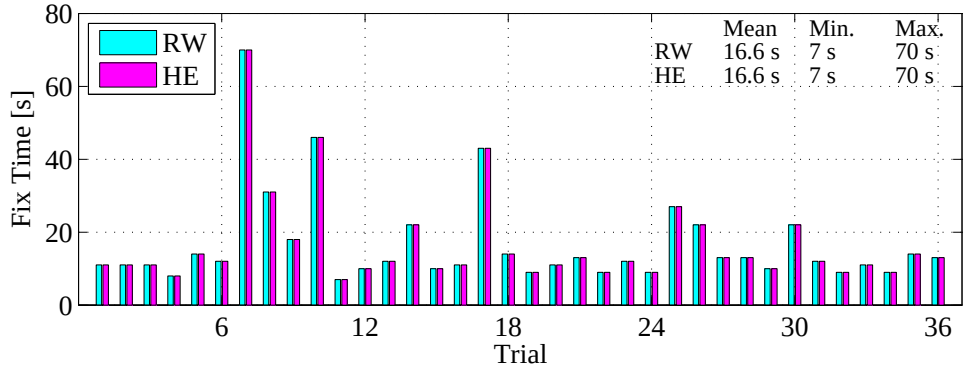


Figure A.13: *Test 2 cold-start ambiguity resolution fix times.*

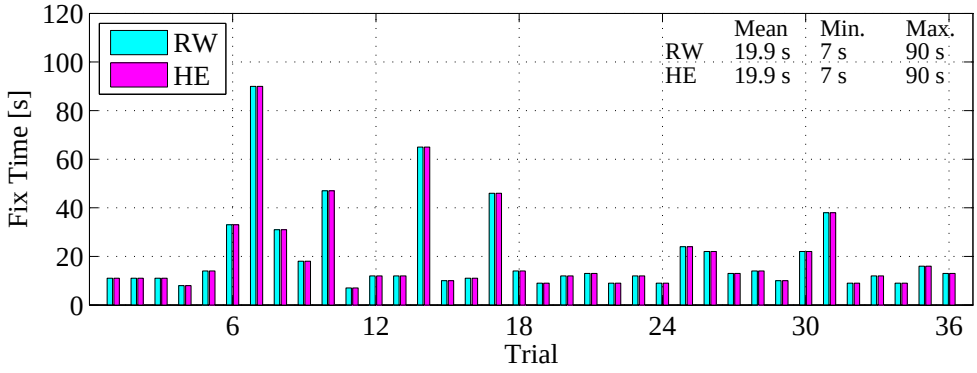


Figure A.14: *Test 3 cold-start ambiguity resolution fix times.*

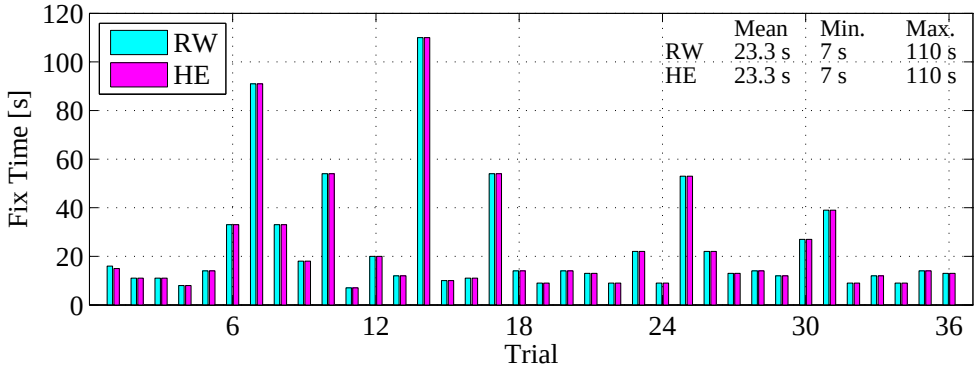


Figure A.15: *Test 4 cold-start ambiguity resolution fix times.*

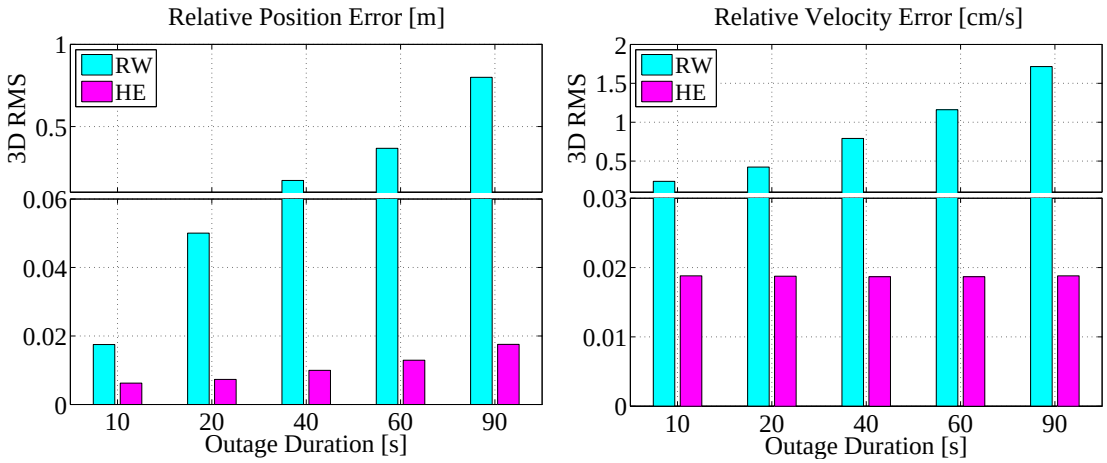


Figure A.16: *Test 2* 3D RMS *errors during signal outage.*

Table A.1: *Test 2 warm-start ambiguity resolution fix time statistics.*

Outage	10 s		20 s		40 s		60 s		90 s	
Model	RW	HE	RW	HE	RW	HE	RW	HE	RW	HE
Mean	4.1	4.0	3.7	3.7	2.3	2.2	4.1	2.5	9.7	3.8
Min.	1	1	1	1	1	1	1	1	3	1
Max.	61	61	55	55	32	31	26	13	48	39

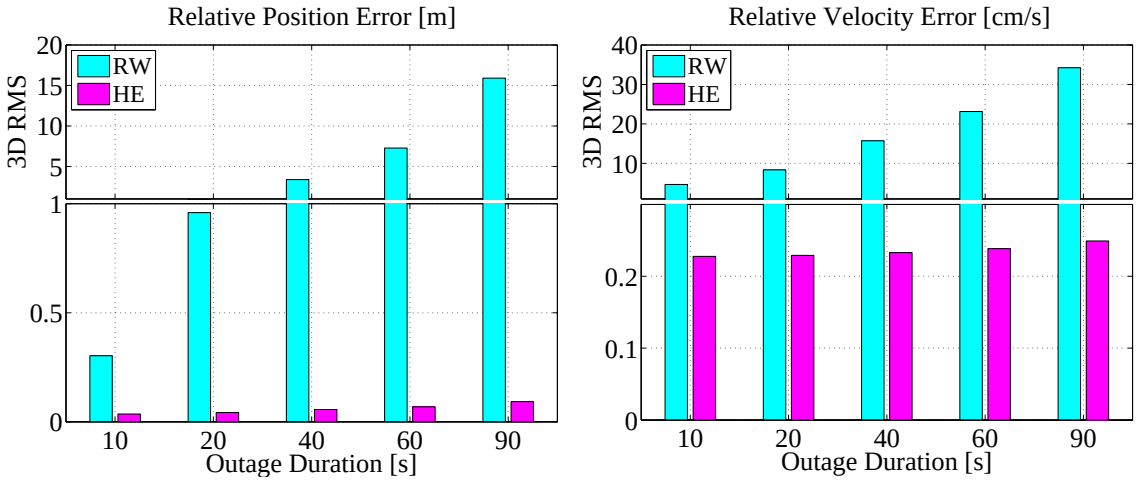


Figure A.17: *Test 6* 3D RMS *errors during signal outage.*

Table A.2: *Test 6 warm-start ambiguity resolution fix time statistics.*

Outage	10 s		20 s		40 s		60 s		90 s	
Model	RW	HE	RW	HE	RW	HE	RW	HE	RW	HE
Mean	170.8	164.3	172.9	165.3	180.8	167.2	188.9	184.8	181.4	155.1
Min.	2	1	2	1	5	1	5	1	6	1
Max.	589	589	579	579	559	559	539	539	509	509