Nonlinear Observer for Tightly Coupled Integrated Inertial Navigation Aided by RTK-GNSS Measurements

Jakob M. Hansen, Tor Arne Johansen, Nadezda Sokolova, and Thor I. Fossen, Fellow, IEEE

Abstract—A modular nonlinear observer is considered for tightly coupled integration of inertial measurements with global satellite measurements. A real-time-kinematic (RTK) approach is utilized where the rover and base station measure pseudoranges, carrier-phase, and carrier-phase derived Doppler, to be used in a dual receiver configuration. The modular observer design consists of a nonlinear attitude observer and a translational motion observer (TMO). The attitude observer represents the vehicle attitude as unit quaternions and estimates the gyro bias. A TMO based on the double-differenced measurements between the rover and base station receiver is proposed. The ambiguities introduced by the carrier-phase measurements are included in the state vector and are initially considered real valued, later to be resolved to integer values. The advantages of the nonlinear observer compared with a Kalman filter are reduced computational load, no linearization requirement of the model, and proven stability. The proposed observer is verified using experimental data from flights with an unmanned aerial vehicle, where the position estimates are shown to be within 2–4 cm of a global positioning system L1-based RTK reference solution.

Index Terms—Global navigation satellite system, inertial navigation, observers, sensor fusion, state estimation, unmanned autonomous vehicles.

I. INTRODUCTION

INTEGRATING measurements from an inertial measurement unit (IMU) with global navigation satellite system (GNSS) measurements is a common method for the estimation of position, linear velocity, and attitude (PVA) of a vehicle. The most common integration scheme is the loosely coupled approach where inertial measurements are integrated with GNSS position and velocity estimates. The position and velocity estimates used in a loosely coupled system are supplied by a GNSS receiver, where the navigation solution is determined by an integration algorithm, most typically a Kalman filter. Due to the unknown tuning of the filter in the GNSS receiver, it can be preferred to use the measurements from the satellites directly in navigation systems to optimize its performance. A tightly coupled integration scheme can be utilized where inertial measurements are integrated with pseudorange, carrier-phase, and Doppler measurements, thereby correcting the state estimates using measurements in the range domain rather than in the position domain. In general, tightly coupled systems have higher performance than loosely coupled systems as a higher level of control of the nuisance and noise terms can be accounted for. A further advantage of the tightly coupled system is that aiding from even a few satellites can be used, whereas for loosely coupled systems, at least four satellites must be available for position determination in the receiver. Moreover, at least four satellites are required for obtaining aiding information in loosely coupled systems, whereas tightly coupled systems can offer partial aiding with fewer satellites. This increases the robustness of tightly coupled solutions to satellite signal obstruction.

Tightly coupled multisensor systems using filters, such as the extended Kalman filter (EKF), have been subject of extensive research (see [1], [2]). In [3], enhancements for a tightly coupled integration on small unmanned aerial vehicles (UAVs) are presented. A low-cost GNSS/IMU integration using a microelectromechanical system (MEMS) IMU with a nonlinear KF approach is developed in [4], while [5] considers a tightly coupled integration using a Kalman filter for accurate mapping using an aerial platform. A comparative study of loosely, tightly, and ultratightly coupled systems was investigated in [6] using a wide range of inertial sensor grades and the GIGET tool.

Higher accuracy can be achieved by taking atmospheric disturbances into account, such as ionospheric and tropospheric delay of the satellite signal. These disturbances can, to some extent, be modeled in nominal conditions, or mitigated by using dual-frequency or multifrequency GNSS receivers. Multifrequency receivers facilitate estimation and compensation of the ionospheric delay (see [2, Sec. 8.6] or [7]). However, for low-cost applications, the high cost of the necessary receiver front ends presently prevents wide applicability of multifrequency systems. Furthermore, multifrequency receivers and antennas often require more space when mounted and can therefore limit their inclusion in constrained applications, such as unmanned aerial vehicles.

Inclusion of the carrier-phase measurement allows for achieving higher accuracy if the introduced integer ambiguities can be resolved. In [8], a low-cost multiple GNSS antenna configuration is utilized to aid in the ambiguity resolution...
performed. Multiple approaches for estimation of the integer ambiguities have been studied, e.g., [9]–[11] proposing the widely used Least-squares AMbiguity Decorrelation Adjustment (LAMBDA) method, or [12] where a computationally efficient method was proposed using a common-position-shift approach. In [13], an approach using an aiding INS to resolve the ambiguities was investigated, while [14] proposes a near-real-time method based on measurements from multiple epochs. Another method is the real-time-kinematic (RTK) solution, where a dual-receiver configuration with a stationary base station and a static or moving rover (i.e., a vehicle) is used. The base station broadcasts its satellite measurements to the rover, which in turn performs differencing of the measurements for use in the navigation system. With a known base station position precise, positioning of the rover is possible. It is also possible to utilize a moving base station: if the base station position is always known, the precision can be retained, and otherwise, inertial sensors might be added to the base station to offer similar state estimation as for the rover. The RTK solution operates under the assumption that the separation between the rover and base station (called the baseline) is short, such that the atmospheric signal delays observed by the two receivers are similar. In [15] and [16], a low-cost open source RTK solution is developed. Some environments can decrease the accuracy of the acquired solution due to satellite blockage, which is the topic of [17], where Doppler measurements are used to aid in the RTK computations. Multipath caused by reflected satellite signals can also cause errors in the pseudorange and carrier-phase measurements, but can to some extent be remedied, e.g., by the proposed method of [18]. In [19], the time differenced carrier-phase measurements are used instead of delta-range measurements in a tightly coupled single-receiver system, showing improvements to velocity and attitude estimation, while introduction of a base station was encouraged to obtain centimeter accuracy. Other approaches include the precise point positioning (PPP) [20] or differential GNSS [21], which can give results with precision similar to RTK systems. In [22], a tightly coupled Kalman filter is proposed for a low-cost sensors solution, where the single-differenced (SD) ambiguities are included in the state vector, and the Kalman filter solution is verified using automotive experimental data.

When low-cost sensors are used, care should be taken to avoid long time periods of satellite obstruction, as the quality of the inertial sensors will introduce fast diverging state estimates. The measurement noise levels will be higher for low-cost sensors, which will introduce large output noise and might prevent correct resolution of the phase ambiguities.

Applications for tightly coupled RTK GNSS/INS integration are many and varied. Here, the focus will be on UAVs where the applications could include automated landing. UAVs can be landed in vertically placed nets at landing sites or on ships, which require high precision. This paper will focus on a low-cost solution for UAV flights close to an area of interest equipped with a stationary base station.

Previously, only loosely coupled GNSS/INS nonlinear observers have been proposed, leaving the EKF-based observers to dominate the field of tight integration. However, recently nonlinear observers have gained interest for tightly coupled GNSS/INS integration [23]–[26], where long baseline configurations have been considered. Nonlinear observers have some advantages compared with Kalman filters, such as proven stability results for nonlinear systems, no need for linearization of the kinematic model, and lower computational load as shown in, e.g., [27] where a tightly coupled nonlinear observer was shown to constitute only around 25% of a comparable multiplicative EKF (MEKP) implementation.

A modular observer structure will be considered here, consisting of the nonlinear attitude estimator proposed in [28] and a translational motion observer (TMO). The observer structure is based on results in [29] and [30].

This paper is an extension of the work in [27] and [31], here proposing a TMO for estimation of position, linear velocity, and integer ambiguities. The injection terms for the observer are the errors between estimated and measured double-differenced (DD) pseudorange, carrier-phase, and carrier-phase derived Doppler measurements. The method can be implemented using low-cost sensors, such as MEMS-based inertial sensors and single-frequency GNSS receivers. This paper offers two main contributions to the previous work of the authors.

1) Time Derivative of Carrier-Phase Measurements: An additional aiding measurement is considered, and the observers have been expanded to include the time derivative of the carrier-phase measurement. This measurement can be utilized for injection terms in the DD observer, and require determination of satellite velocity. The time derivative of the carrier-phase measurement is used rather than the raw Doppler measurement, due to the lower measurement noise.

2) Experimental Verification: The proposed observer has been experimentally verified by data acquired using a fixed-wing UAV equipped with a low-cost sensor suite consisting of an ADIS-16488 IMU and a uBlox LEA-6T GPS L1 receiver. A second uBlox receiver of the same model was used as the base station. The reference was determined with an open source RTK solution using the same receiver pair.

A. Notation and Preliminaries

A column vector $x \in \mathbb{R}^3$ is denoted $x := [x_1; x_2; x_3]$ with transpose $x^T$ and vector norm $\|x\|_2$. The skew-symmetric matrix of a vector $x$ is given as

$$S(x) := \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}.$$  

A unit quaternion, $q = [q_r; q_i]$, is described as consisting of a real part, $q_r \in \mathbb{R}$, and a vector part, $q_i \in \mathbb{R}^3$, where $\|q\|_2 = 1$. A vector $x \in \mathbb{R}^3$ can be represented as a quaternion with zero real part, $\mathbf{x} = [0; x]$. The product of two quaternions $q_1$ and $q_2$ is the Hamiltonian product denoted $q_1 \otimes q_2$.

Superscripts will be used to signify which coordinate frame a vector is decomposed in. Rotation between the frames may be represented by a quaternion, $q_{ot}^r$, describing the rotation
from coordinate frame $a$ to $c$, with the corresponding rotation matrix $R(a^c \in SO(3)$. Various coordinate frames will be used, where the earth-centered-earth-fixed (ECEF) frame is denoted with $e$, while $b$ will be used for Body-frame with $n$ denoting the north-east-down frame. The earth-centered-inertial (ECI) frame will be denoted with $i$ and will only be used for angular rates, such as the Earth rotation rate $\omega_{ei}$, as the $i$ will later be used to signify signals from the $i$th satellite.

II. Problem Formulation

A dual-receiver configuration consisting of a moving rover and a stationary base station, each equipped with a GNSS receiver, is considered. The base station measures pseudorange, carrier-phase, and carrier-phase derived Doppler and transmits these to the rover for processing, followed by position and velocity estimation. The rover will, in addition, be equipped with inertial sensors providing specific force and angular rate measurements.

The objective is to estimate PVA of a moving vehicle by tightly coupled integration of mentioned measurements. Estimation of the rover PVA as well as the baseline, i.e., the vector (displacement) between rover and base station, is of interest.

The kinematic strapdown equations of the rover are given as

$$
\dot{p}_c^e = v_c^e \quad (1)
$$

$$
\dot{v}_c^e = -2S(o_{ie})v_c^e + f_c^e + g_c^e(p_c^e) \quad (2)
$$

$$
\dot{q}_b^e = \frac{1}{2} q_b^e \otimes m_b^e - \frac{1}{2} m_{ie}^e \otimes q_b^e \quad (3)
$$

where the PVA of the rover is described by $p_c^e$, $v_c^e$, and $q_b^e$, respectively. The Earth rotation rate in ECEF with respect to ECI decomposed in the ECEF frame is denoted $o_{ie}$, and is a known constant. The effect of the Earth rotation will be very small compared with the other terms of (2), and can often be disregarded during implementation. However, here it will be included for completeness. The rotation rate $o_{ib}^e$ describes the body rotation rate with respect to the ECI frame. The specific force is given by $f_c^e$, while $g_c^e(p_c^e)$ specifies the position-dependent plumb-bob gravity vector. The base station is assumed to be stationary, $\dot{p}_s^e = 0$.

A. Measurement Assumptions

It is assumed that the rover is equipped with an IMU, a GNSS receiver, and a magnetometer. The following assumptions on the measurements are made.

**Assumption 1:** Specific force measurements are available as $f_{IMU}^b = f^b$.

**Remark 1:** The accelerometer bias estimation module in [32] and [33] can be added as a first stage providing a bias-compensated acceleration measurement in cascade with the proposed observer. Since this is a cascade, this module has no impact on stability of the error dynamics.

**Assumption 2:** The rotational rate of a gyro is measured as $\omega_{ib,IMU}^e = \omega_{ib}^e + b^b$. The gyro bias $b^b \in \mathbb{R}^3$ is assumed to be slowly time-varying, $\dot{b}^b = 0$, satisfying $\|b^b\|_2 \leq M_b$ for a known bound $M_b$.

**Assumption 3:** A magnetometer measures the magnetic field vector at the rover, $m_{MAG}^b = m^b$, while the position-dependent Earth magnetic field $m_e$ in ECEF is assumed known.

Furthermore, the rover and the base station are assumed to be equipped with GNSS-receivers measuring pseudorange, carrier-phase, and carrier-phase derived Doppler, from at least four common satellites, $m \geq 4$. The satellite positions and velocities are assumed known, which can be satisfied by determination using the broadcast ephemeris data.

The pseudorange and carrier-phase measurements at a location $p^e_i$ transmitted by the $i$th satellite [see Fig. 1(a)] can be generalized as

$$
\rho_i = \psi_i + \beta + \gamma_i + \epsilon_\rho + n_{\rho, i} \quad (4)
$$

$$
\lambda \phi_i = \psi_i + \lambda N_i + \beta + \gamma_i + \epsilon_\phi + n_{\phi, i} \quad (5)
$$

where the pseudorange and carrier-phase measurements from the $i$th satellite are denoted by $\rho_i$ and $\phi_i$, for $i = 1, 2, \ldots, m$. The geometric range between receiver and satellite is denoted $\psi_i := \|p_i^e - p_i^r\|_2$, where $p_i^r$ is the satellite position. The carrier-phase integer ambiguity is denoted $N_i$ with signal wavelength $\lambda$. The parameters $\epsilon_\rho$ and $\epsilon_\phi$ denote the combined orbital, ionospheric, and tropospheric delays of the satellite measurements. Furthermore, $\epsilon_\phi$ includes antenna-induced errors. The $n_i$ parameters cover individual stochastic errors, such as measurement noise and multipath effect. The noise terms have been explicitly included for the aiding measurements to ensure that the measurement noise is not considered canceled in the differences introduced later. The measurement noise of the aiding measurement will change magnitude when differenced which should be taken into account in the tuning of the observer. The noise terms for the inertial measurements will not be explicitly included as they will not undergo operations that will change their characteristics. All noise sources will be systematically considered in the tuning of the observer. The satellite clock error $\gamma_i$ can be corrected by applying the clock correction coefficients from the navigation data message. The receiver clock bias, describing the timing difference between receiver and system time, affects all coinciding range measurements similarly (see [2, Sec. 8.4.3]). In the following, the clock range bias is treated as a common mode bias, $\beta \in \mathbb{R}$, and is assumed to be slowly time varying.

The bias is expressed as $\beta = c \Delta_{r,i}$, where $\Delta_{r,i}$ is the receiver clock bias and $c$ is the speed of light. Similarly, the satellite clock error can be expressed as $\gamma_i = c \Delta_{s,i}$, where $\Delta_{s,i}$ is the bias of the $i$th satellite clock.

Due to large measurement noise, the raw Doppler measurements are usually not included as aiding information. The measurements can either be filtered, which has the drawback of reducing the sample frequency, removing high frequency motion and introduce phase lag, or a substitute can be determined as the time derivative of the carrier-phase measurements. In [34], the carrier-phase derived Doppler is obtained by use of a finite impulse response filter, resulting in accurate velocity measurements for a standalone global positioning system (GPS) receiver. In the following, the time
derivative of the carrier phase will be considered [35]:
\[ \dot{\phi}_i = \psi_i + \dot{\beta} + \dot{\gamma}_i + n_{\psi,i} \]
where the ambiguity and multipath terms have been eliminated due to the short time horizon of the derivative. Furthermore, by assuming constant atmospheric delays over the time interval between consecutive phase measurements, these can also be eliminated, where any residual delays will be absorbed by the noise term [35].

A generalized model for the time derivative of the phase measurement can be expressed as
\[ \dot{\nu}_i = h_i^T (v^e_i - \nu^e_i) + \dot{\beta} + \dot{\gamma}_i + n_{\nu,i} \]
where the satellite velocity is denoted as \( v^e_i \), and the line-of-sight vector is \( h_i = (p^e_i - p^e_j)/\psi_i \). The constructed carrier-phase derived measurement is denoted as \( \nu_i \), while \( n_{\nu,i} \) describes the stochastic errors.

Remark 2: The carrier-phase measurements \( \varphi \) are given in units of cycles, and can be converted to meters by multiplication with the wavelength. The carrier-phase derived Doppler measurements \( \nu \) are given in Hertz and can be converted to velocity described in meters per second by multiplication with the wavelength.

Superscripts of \( r \) and \( s \) will be used to distinguish between the measurements at rover and base station, respectively. These superscripts are not to be confused with the coordinate frame as the satellite measurements will always be given in the ECEF-frame. It is assumed that the rover is sufficiently close to the base station, such that the ionospheric and tropospheric delays are spatially correlated between the two receivers and will therefore cancel, i.e., \( \varepsilon^e_r = \varepsilon^e_s, \varepsilon^o_r = \varepsilon^o_s \). While it is possible to obtain centimeter accuracy of the position estimates with the baseline of 20 km for resolved ambiguities (see [2, Sec. 10.2]) maintaining high accuracy requires shorter baselines. Short baselines, less than 10 km, are preferred to reduce cycle slips, while for longer baselines, the accuracy degrades due to decorrelation of the atmospheric propagation errors. This can be mitigated by use of dual-frequency GNSS measurements and modeling of the troposphere. The atmospheric disturbances can, therefore, be corrected for by differencing measurements at rover position with measurements at the base station. The GNSS measurement configuration is depicted in Fig. 1(b) showing two satellites, \( i \) and \( j \), and two receivers, \( r \) and \( b \), where the baseline is shown as a dashed line. It is vital when differencing measurements that they are acquired at the same time epoch.

The SD measurements between rover and base station are given as
\[ \Delta \rho_i = \Delta \psi_i + \Delta \beta + n_{\Delta \rho,i} \]
\[ \dot{\lambda} \Delta \rho_i = \dot{\lambda} \Delta \psi_i + \dot{\lambda} \Delta N_i \dot{\lambda} + \dot{\lambda} \Delta \beta + n_{\Delta \rho,i} \]
\[ \dot{\lambda} \Delta \nu_i = h_i^T (v^e_i - \nu^e_i) + \dot{\beta}_r + \dot{\gamma}_i + n_{\Delta \nu,i} \]
where \( \Delta \rho_i := \rho^e_i - \rho^e_j, \Delta \psi_i := \psi^e_i - \psi^e_j, \Delta \nu_i := \nu^e_i - \nu^e_j, \Delta N_i := N_i^e - N_j^e, \Delta \beta := \beta^r - \beta^s, \) and \( \Delta \psi_i := \psi^e_i - \psi^e_s \) is the geometric baseline between rover and base station. The differences noise terms are given as; \( n_{\Delta \rho,i} := n_{\rho,i} - n_{\rho,j}, n_{\Delta \psi,i} := n_{\psi,i} - n_{\psi,j}, n_{\Delta \nu,i} := n_{\nu,i} - n_{\nu,j}. \) The normalized line-of-sight vectors \( h_i^e \) and \( h_i^s \) between the \( i \)th satellite and rover and base station are given as
\[ h_i^e = \frac{p^e_i - p^e_j}{\| p^e_i - p^e_j \|_2}, \quad h_i^s = \frac{p^e_i - p^e_s}{\| p^e_i - p^e_s \|_2} \]
From (8) and (9), it is evident that \( \Delta \beta \) is the same for all satellites so it can be canceled by further differencing the measurements this time between satellites. This is typically achieved by selecting the satellite with highest elevation as reference. The DD measurement principle is depicted in Fig. 1(c), while the measurement model is given as
\[ \nabla \Delta \rho_{ij} = \nabla \Delta \psi_{ij} + n_{\nabla \Delta \rho,ij} \]
\[ \dot{\lambda} \nabla \Delta \psi_{ij} = \dot{\lambda} \nabla \Delta \psi_{ij} + \nabla \Delta N_{ij} \dot{\lambda} + n_{\lambda \nabla \Delta \psi,ij} \]
\[ \dot{\lambda} \nabla \Delta \nu_{ij} = h_j^T (v^e_j - \nu^e_j) - h_i^T (v^e_i - \nu^e_i) + h_j^T v^e_j \]
\[ - h_i^T v^e_i + n_{\lambda \nabla \Delta \nu,ij} \]
where \( \nabla \Delta \rho_{ij} := \Delta \rho_j - \Delta \rho_i, \nabla \Delta \psi_{ij} := \Delta \psi_j - \Delta \psi_i, \nabla \Delta \nu_{ij} := \Delta \nu_j - \Delta \nu_i, \nabla \Delta \psi_{ij} := \Delta \psi_j - \Delta \psi_i, \nabla \Delta N_{ij} := \Delta N_j - \Delta N_i \) with the \( j \)th satellite representing the reference satellite. The measurement noise is \( n_{\nabla \Delta \rho,ij} := n_{\Delta \rho,j} - n_{\Delta \rho,i}, n_{\nabla \Delta \psi,ij} := n_{\Delta \psi,j} - n_{\Delta \psi,i}, n_{\nabla \Delta \nu,ij} := n_{\Delta \nu,j} - n_{\Delta \nu,i}. \)

Due to measurement differencing between satellites, the number of available observations will be \( m - 1 \), where \( m \) is the number of tracked satellites. The DD measurements will
be used in injection terms of the proposed observer structure. The SD measurements are not utilized in the observer structure due to the presence of the receiver clock bias.

When differencing GNSS measurements, a rule of thumb is that the stochastic error will increase in standard deviation (STD) by $\sqrt{2}$, while the systematic errors will decrease (see [36]).

### III. Noise Analysis

A common assumption of the EKF is that the measurement and process noise are Gaussian. The nonlinear observer structure proposed here does not make any specific assumptions on the distribution of GNSS measurement noise in (4)–(10) and (12)–(14), or the inertial measurement noise introduced with the specific force, angular rate, and magnetometer measurements. However, when tuning the observer gains, it is advantageous to have the knowledge of the noise levels and types. It is, therefore, desired to investigate the noise distribution and determine the STD of the undifferenced (UD) and DD GNSS measurement noise. The noise type is the same for all types and brands of receivers, whereas the STDs are receiver and antenna specific. For the inertial measurement noise, industrial data sheets offer accurate noise characteristics. If the noise characteristics cannot be assumed Gaussian white, the state-space model can be augmented to include colored noise models with white driving noise (see [2, Sec. 4.6.3] or [37, Sec. 7.5]). Section IV will, therefore, only consider the GNSS measurement noise.

Experimental data were collected at 1 Hz, by two identical u-Blox LEA-6T receivers with a baseline of 1 m. The experiment was carried out at latitude 63.4° and longitude 10.5°, Norway, on a rooftop. Some multipath effect is experienced, which affect the UD measurements, while it can be mostly mitigated for the DD measurements due to the short baseline. The positions of the receivers were determined by averaging over a 24-h-long data set. Some satellites were occasionally obstructed and the noise levels of the GNSS measurements are, therefore, found during a time interval with multiple common satellites visible by both receivers. The satellites are commonly visible for 3–6 h during good open sky conditions.

In the case of the UD measurements, the raw pseudorange, carrier-phase, and carrier-phase derived Doppler measurements are corrected for: time of transmission (the satellite moves during the signal transmission time), atmospheric disturbances, and receiver clock bias (estimated using a Kalman filter). Furthermore, the geometric range from satellite to receiver $r'_i$ and $\psi'_i$ is computed. The tropospheric delay can be modeled using, e.g., [38], whereas the ionospheric delay might be modeled using the Klobuchar model. For the UD noise terms, some residual errors will be present due to the imperfection of the applied models. Furthermore, uncertainty of the estimated receiver clock and multipath effects remains.

For the DD measurements, the noise terms are found by isolation of $n_i$ in (12)–(14). Given the short baseline (1 meter), the ionospheric and tropospheric delays are considered to be completely removed, whereas the residual atmospheric effects are negligible, such that the noise term only consists of receiver noise and multipath.

The noise of the DD measurements is shown in Fig. 2(a)–(c), with six available satellites. The noise estimates resulting in five DD measurements. The noise estimates have been detrended to remove first-order dependences and initial offset, which is common practice, e.g., when determining scintillation effects (see [39] or [40]). The colors indicate the measurements from different satellites. The average STDs of the estimated measurement noise are presented in Table I, with the noise of the UD measurements for comparison. The STDs correspond well with the results found in [1, Sec. 9.4.2.4] for an update rate of 1 Hz. Moreover, according to [1, Sec. 9.4.2.4], the STDs might be increased with higher receiver sampling rates.

To obtain a large time window for noise analysis, some of the satellites have low elevation at the edges of the time window, and this results in an increase in STD, as seen in, e.g., the yellow and red graphs in Fig. 2(c) around 9000–10 000 s. Nevertheless, these satellites have been included to estimate an average STD of the noise terms under various conditions. Similar effects can be seen in [41], where a filter for the separation of the multipath components is proposed, while also locating reflecting objects. Short and long baselines are considered in [42] and [43], where UD, SD, and DD GNSS signals from several constellations are analyzed.

The DD pseudorange and carrier-phase derived Doppler measurements appear almost white, whereas the carrier-phase measurements include nonwhite behavior. There are two common practices for handling this in observer structures: 1) augment the observer to model states for the nonwhite noise processes or b) assume a white distribution with larger STD than the actual distribution to overbound the measurement noise. The overbounding method was proposed in [44], and used, e.g., in [45] where multipath, receiver noise, and atmospheric delays were investigated for pseudoranges in an SD configuration. In general, the code and phase noise can be represented by overbounding white noise, even though a Gauss–Markov process is a more suitable model. It might be beneficial to use slightly inflated noise characteristics even when employing Gauss–Markov processes. Additionally, the distributions cannot be expected to be stationary as several factors can cause changes on the fly, such as multipath (especially at the rover), elevation angle, rover dynamics, and $C/N_0$ levels. In conclusion, it might be advantageous to use higher values for the STDs when used for observer tuning to accommodate for imperfect Gaussian white noise behavior.

### IV. Observer Structure

In Sections V and VI, a nonlinear observer is proposed based on DD GPS $L_1$ measurements. A modular observer
structure consisting of a nonlinear attitude estimator and a TMO is considered. The framework is the same as presented in [27] and [31] although a different TMO is used: here, the dual-receiver configuration is used leading to differenced measurements, whereas in [27], a single receiver was used to obtain a tightly coupled navigation solution.

The observer structure, shown in Fig. 3, consists of two GNSS receivers, IMU, magnetometer (MAG), nonlinear attitude observer, a TMO, and a Riccati solver with a gain estimator. Two feedback loops are present in the structure: the feedback of the estimated rover position and velocity to the computation of gain matrices, and an additional feedback of the estimated rover position and velocity to the TMO to the attitude observer, from the TMO to the attitude observer, and additionally the gyro bias is determined. The nonlinear attitude observer proposed in [29], [32], and [47] is used, where the carrier-phase derived Doppler measurement is based on the carrier-phase change, and expresses the change in geometric distance and clock drift over a time interval (see [2, Sec. C.4.2]). The carrier-phase derived Doppler measurement, therefore, needs to be compensated for satellite motions during the time interval when compared with \( v^f_i \). Several options are available for estimating the average satellite velocity of the interval. Here, \( v^f_i \) will express the satellite velocity averaged over the beginning and end of the relevant GNSS receiver sample interval.

The “Riccati Solver & Gain Estimator” block considers the choice of gain matrix \( K \) for the TMO. In order to compute the injection terms, the estimated position and velocity of the rover and base station is fed back from the TMO.

When considering the stability of the observer structure, it is advantageous to examine the individual parts of the modular structure. The nonlinear attitude observer was proven to be semiglobally exponentially stable (SGES) with respect to attitude initialization in [28]. Following the results of [27], the equilibrium point of the DD observer proposed in Section VI is exponentially stable. The complete observer structure will be exponentially stable with a semiglobally region of attraction with respect to attitude initialization due to the feedback interconnection and the SGES properties of the attitude estimator. A rigorous stability proof supporting this is given in [46]. The observer position and velocity estimates can be accurately initialized with the procedure proposed in [27].

In the following, the nonlinear attitude observer will be introduced in Section V, while the TMO will be introduced utilizing the DD receiver configuration in Section VI.

V. NONLINEAR ATTITUDE OBSERVER

Section VI introduces the nonlinear attitude observer. The orientation of the stationary base station is unimportant, and only the attitude of the moving rover will be considered.

The attitude of the rover, given as the rotation between Body- and ECEF-frame, is represented as a unit quaternion to avoid singularities at steep pitch angles. A recent attitude observer proposed in [29], [32], and [47] is used, where additionally the gyro bias is determined. The nonlinear attitude observer with the injection term \( \dot{\sigma} \) is introduced as

\[
\dot{\hat{q}}_b^c = \frac{1}{2} \hat{q}_b^c \otimes (\overline{\mathbf{ω}}^c_{\text{IMU}} - \bar{b} + \sigma) - \frac{1}{2} \overline{\mathbf{ω}}^c_e \otimes \hat{q}_b^c \tag{15}
\]

\[
\dot{\bar{b}}^h = \text{Proj}(-k_1 \dot{\sigma}, \|\bar{b}^h\|_2 \leq M_{\bar{b}}) \tag{16}
\]

\[
\dot{\sigma} = k_1 \sqrt{\bar{b}^h} \times R(\hat{q}_b^c)^T \overline{\mathbf{ω}}^c_1 + k_2 \bar{b}^h \times R(\hat{q}_b^c)^T \overline{\mathbf{ω}}^c_2 \tag{17}
\]

where \( \text{Proj}(\cdot, \cdot) \) describes the projection operator limiting the gyro bias estimate to a sphere with radius \( M_{\bar{b}} \) with \( M_{\bar{b}} > M_{\bar{b}} \).

The tuning constants \( k_1, k_2 \), and \( k_1 \) are chosen positive and sufficiently large.
The injection term considers two vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ in the Body-frame and their corresponding vectors $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{v}}_i$ in the ECEF-frame. These vectors can be chosen in various ways, but will here be based on specific force and magnetic field

$$\hat{\mathbf{u}}_i = \frac{f_{\text{IMU}}^i}{\|f_{\text{IMU}}^i\|_2}, \quad \hat{\mathbf{v}}_i = \frac{m_{\text{MAG}}^i}{\|m_{\text{MAG}}^i\|_2} \times \hat{\mathbf{u}}_i$$

where $m^i$ is the position-dependent magnetic field vector of the Earth assumed known for the rover’s estimated position and the specific force estimate $\hat{f}^e$ is supplied by the TMO.

An advantage of this attitude observer is the constant gains of the injection term, leading to a lower computational load compared with other estimators, where a Kalman filter approach will have a significantly higher load due to the matrix inverse and the additional propagation of a covariance matrix. In [27], a similar nonlinear observer is compared with an MEKF, where the attitude observer was seen to constitute only around 20% of the computational load of the MEKF attitude part (and around 25% when the entire observer structures were compared).

VI. DOUBLE-DIFFERENCED NONLINEAR OBSERVER

Denoting the $i$th satellite as the reference satellite, the DD measurement errors can be used as injection terms in the proposed TMO

$$\hat{\mathbf{u}}_i^c = \hat{\mathbf{u}}_i + \sum_{j=1}^{m-1} \left( K_{ij}^{pp} e_{\rho,ij} + K_{ij}^{pp} e_{\phi,ij} + K_{ij}^{pu} e_{o,ij} \right)$$

$$\hat{\mathbf{v}}_i^c = -2S(\sigma_{\rho}) \hat{\mathbf{v}}_i + \hat{f}^e + g^e(\hat{\mathbf{u}}_i)$$

$$\hat{\mathbf{v}}_i = -R(\hat{\mathbf{v}}_i^c) S(\hat{\mathbf{v}}_i)^T_{\text{IMU}}$$

$$\hat{\mathbf{v}}_i = \sum_{j=1}^{m-1} \left( K_{ij}^{pp} e_{\rho,ij} + K_{ij}^{pp} e_{\phi,ij} + K_{ij}^{pu} e_{o,ij} \right)$$

$$\hat{f}^e = R(\hat{\mathbf{v}}_i^c) f_{\text{IMU}} + \hat{\mathbf{v}}_i$$

$$\hat{\mathbf{v}}_i^c = \sum_{j=1}^{m-1} \left( K_{ij}^{np} e_{\rho,ij} + K_{ij}^{np} e_{\phi,ij} + K_{ij}^{nu} e_{o,ij} \right)$$

$$\nabla \Delta \hat{N} = \sum_{j=1}^{m-1} \left( K_{ij}^{np} e_{\rho,ij} + K_{ij}^{np} e_{\phi,ij} + K_{ij}^{nu} e_{o,ij} \right).$$

The gains to be determined, $K^{**}$, are considered slowly time varying. Here, the auxiliary term $\xi$ is presented in order to determine the estimated specific force in ECEF frame [see (21)]. The advantage of using DD measurements is that the receiver clock bias is canceled and can therefore be excluded from estimation, while in single-receiver configurations, it is included in the state vector (see [46]).

The DD phase ambiguities $\nabla \Delta \hat{N}$ are represented as a vector with $m - 1$ elements described as the difference between the SD ambiguities of the $i$th and $j$th satellite, i.e., $\nabla \Delta \hat{N} = [\Delta \hat{N}_1 - \Delta \hat{N}_i; \Delta \hat{N}_2 - \Delta \hat{N}_i; \ldots; \Delta \hat{N}_j - \Delta \hat{N}_i]$. The ambiguity estimates are propagated by the TMO as real-valued estimates and can after some convergence be fixed to integer values, which will be discussed in detail in Section VI-B.

It is desired to include the base station position in the observer estimation, such that the observer does not rely on access to a presurveyed position, but can be initialized with a more inaccurate estimate of the base station position.

The structure of the TMO is similar to [31] with additional injection terms based on the carrier-phase derived Doppler measurements. The injection terms are the difference between measured and estimated DD satellite signals:

$$e_{\rho,ij} := \nabla \Delta \rho_{ij} - \nabla \Delta \phi_{ij}, \quad e_{\phi,ij} := \lambda(\nabla \Delta \phi_{ij} - \nabla \hat{\phi}_{ij})$$

The gains to be determined, $K^{**}$, are considered slowly time varying. Here, the auxiliary term $\xi$ is presented in order to determine the estimated specific force in ECEF frame [see (21)]. The advantage of using DD measurements is that the receiver clock bias is canceled and can therefore be excluded from estimation, while in single-receiver configurations, it is included in the state vector (see [46]).

The DD phase ambiguities $\nabla \Delta \hat{N}$ are represented as a vector with $m - 1$ elements described as the difference between the SD ambiguities of the $i$th and $j$th satellite, i.e., $\nabla \Delta \hat{N} = [\Delta \hat{N}_1 - \Delta \hat{N}_i; \Delta \hat{N}_2 - \Delta \hat{N}_i; \ldots; \Delta \hat{N}_j - \Delta \hat{N}_i]$. The ambiguity estimates are propagated by the TMO as real-valued estimates and can after some convergence be fixed to integer values, which will be discussed in detail in Section VI-B.

It is desired to include the base station position in the observer estimation, such that the observer does not rely on access to a presurveyed position, but can be initialized with a more inaccurate estimate of the base station position.

The structure of the TMO is similar to [31] with additional injection terms based on the carrier-phase derived Doppler measurements. The injection terms are the difference between measured and estimated DD satellite signals:

$$e_{\rho,ij} := \nabla \Delta \rho_{ij} - \nabla \Delta \phi_{ij}, \quad e_{\phi,ij} := \lambda(\nabla \Delta \phi_{ij} - \nabla \hat{\phi}_{ij})$$

and the estimated DD geometric baseline is

$$\nabla \Delta \hat{\psi}_{ij} = \|\hat{\rho}_{ij} - \hat{\rho}_{ij}^c\|_2 - \|\hat{\rho}_{ij} - \hat{\rho}_{ij}^c\|_2 + \|\hat{\rho}_{ij} - \hat{\rho}_{ij}^c\|_2$$

Investigating the observer stability, the error states are introduced as $\tilde{\rho}_{ij} := \hat{\rho}_{ij} - \hat{\rho}_{ij}^c, \tilde{\psi}_{ij} := \psi_{ij} - \psi_{ij}^c, \hat{f} := \hat{f}^e - \hat{f}^c, \hat{\psi}_{ij} := \hat{\psi}_{ij} - \hat{\psi}_{ij}^c$, and $\nabla \Delta \hat{N} := \nabla \Delta \hat{N} - \nabla \Delta \hat{N}$, and the state vector of the error dynamics is defined as $x = [\tilde{\rho}_{ij}; \tilde{\psi}_{ij}; \hat{f}; \hat{\psi}_{ij}; \nabla \Delta \hat{N}]$. Here, a combination of (20) and (21) is used to substitute the state $\xi$ with $\hat{f}^c$, as done in [46].

In order to determine the gains $K^{**}$, the injection terms are linearized with respect to the observer states (see Appendix A)

$$e_{\rho,ij} = C_{\rho,ij} \hat{\xi} + \eta_{\rho,ij}$$

$$e_{\phi,ij} = C_{\phi,ij} \hat{\xi} + \eta_{\phi,ij}$$

$$e_{v,ij} = C_{v,ij} \hat{\xi} + \eta_{v,ij},$$

where $\eta_{\cdot,ij}$ consists of measurement noise and higher order nonlinear terms, which can be disregarded in the gain selection under the assumption that the TMO is accurately initialized. The row vectors consist of the following elements: $C_{\rho,ij} = [\hat{f}_{ij}^r, 0, 0, -\hat{b}_{ij}^r, 0], C_{\phi,ij} = [\hat{h}_{ij}^r, 0, 0, -\hat{h}_{ij}^s, \lambda_i,l_{m-1}],$ and $C_{v,ij} = [\hat{f}_{ij}^v, \hat{h}_{ij}^v, 0, \hat{h}_{ij}^v, 0], l_{i,m-1}$ is a row of $m - 1$ zeros with a 1 as the $i$th element. The differenced line-of-sight vectors are described by

$$\hat{h}_{ij}^r = \frac{\hat{\rho}_{ij} - \hat{\rho}_{ij}^c}{\|\hat{\rho}_{ij} - \hat{\rho}_{ij}^c\|_2} - \frac{\hat{\rho}_{ij} - \hat{\rho}_{ij}^c}{\|\hat{\rho}_{ij} - \hat{\rho}_{ij}^c\|_2}$$

$$\hat{h}_{ij}^v = \frac{\hat{\rho}_{ij} - \hat{\rho}_{ij}^c}{\|\hat{\rho}_{ij} - \hat{\rho}_{ij}^c\|_2} - \frac{\hat{\rho}_{ij} - \hat{\rho}_{ij}^c}{\|\hat{\rho}_{ij} - \hat{\rho}_{ij}^c\|_2}.$$
A time-varying measurement matrix consisting of $3(m-1)$ rows can be defined as $C := [C_{p,i}^1; \ldots; C_{p,i}^{m-1}; C_{r,i}^1; \ldots; C_{r,i}^{m-1}; C_{v,i}^1; \ldots; C_{v,i}^{m-1}]$. The $C$ matrix is slowly time varying as the relative motion of the satellites with respect to the receivers is small due to the large separation.

The error dynamics can be determined as

$$\dot{x} = (A - KC)x + \theta_1(t, x) + \theta_2(t, x) + \theta_3(t, x).$$

(35)

Here, $\dot{x}$ is the combined error variable $\tilde{x} := [\tilde{r}; \tilde{b}]$ consisting of the vector part of the quaternion, $\tilde{r} = r - \bar{r}$, and the gyro bias error. The perturbation terms are described as [28]:

$$\theta_1(t, x) := [0; -2S(o_{e,1})x_2 + (g^2(p_f^2) - g^2(p_f^2 - x_1)); 0; 0; 0; 0]$$

and $\theta_2(t, x) := [0; \bar{d}; 0; 0; 0]$, where

$$\bar{d} = (I - R(\tilde{q}^T))R(q_b^e)(S(o_{b,IMU}^b)J_{IMU}^b + J_f^b) - S(o_{e,1})(I - R(\tilde{q}^T))R(q_b^e)J_{IMU}^b - R(\tilde{q}^T)R(q_b^e)P_{(b)}^b J_{IMU}^b.$$

In [28], it is shown that $\|\theta_2(t, x)\| \leq \gamma_3\|\tilde{x}\|_2$, for some positive $\gamma_3$. As in [31], the last perturbation term is a result of the injection term linearization $\theta_3(t, x) := K\eta(t, x)$, where $\eta(t, x) := [\eta_{p,i}; \ldots; \eta_{p,m}; \eta_{r,i}; \ldots; \eta_{r,m}; \eta_{v,i}; \ldots; \eta_{v,m}]$.

Remark 4: In [28] and [46], which this paper is based on, a time-scaling factor $\theta \geq 1$ is introduced in the stability proof. The stability results require the tuning parameter $\theta$ to be chosen sufficiently high, and [28] even offers conservation bounds. However, it is often advantageous to tune the parameter by trial-and-error, increasing the parameter from its lower bound (see also [48]). It is not necessary to include the scaling factor here, as the stability follows from the results of the previous papers. Furthermore, the scaling can be chosen at its lower bound, in the implementation, thereby removing its effect on the observer structure.

The gain matrices should be chosen, such that the nominal linear time-varying closed-loop dynamics $(A - KC)$ is stable (see Section VI-A), where

$$A = \begin{bmatrix} 0 & I_3 & 0 & 0 \\ 0 & 0 & I_3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$K = \begin{bmatrix} K_{pp}^{m-1} & \cdots & K_{pp}^0 & K_{pp}^0 & \cdots & K_{pp}^{m-1} \\ K_{rp}^0 & \cdots & K_{rp}^0 & K_{rp}^0 & \cdots & K_{rp}^0 \\ K_{vp}^0 & \cdots & K_{vp}^0 & K_{vp}^0 & \cdots & K_{vp}^0 \\ K_{np}^0 & \cdots & K_{np}^0 & K_{np}^0 & \cdots & K_{np}^0 \\ K_{1p}^0 & \cdots & K_{1p}^0 & K_{1p}^0 & \cdots & K_{1p}^0 \\ K_{1r}^0 & \cdots & K_{1r}^0 & K_{1r}^0 & \cdots & K_{1r}^0 \\ K_{1v}^0 & \cdots & K_{1v}^0 & K_{1v}^0 & \cdots & K_{1v}^0 \\ K_{1n}^0 & \cdots & K_{1n}^0 & K_{1n}^0 & \cdots & K_{1n}^0 \end{bmatrix}.$$

(36)

The observer system (18)–(23) has an SGES region of attraction with respect to attitude initialization errors and local exponential stability with respect to TMO initialization errors, following the results of [46].

Selection of the reference satellite has a great impact on the performance of the observer, as errors in the measurements from the reference satellite are propagated to all DD measurements. The reference satellite is, therefore, often chosen as the satellite with the highest elevation, since propagation and multipath errors will generally be the smallest (see [49, Sec. 7.3.4]). Another advantage of choosing the satellite with the highest elevation is that this satellite is less likely to be obstructed, thereby keeping phase lock with the receiver. If the chosen reference satellite is obstructed, a new reference has to be selected in order to construct new DD measurements.

A. Gain Selection

Section VI-A will investigate the crucial considerations of gain selection in connection to implementation of the observer. The notation of this section will, therefore, consider discrete time, instead of continuous time used to present the observer structure in Sections IV–VI. The observer should be implemented using discrete time, where the sampling can be dictated by the fastest sensor, usually the IMU, to ensure convergence and stability. For details on discretization and implementation of a similar modular observer structure, see [48].

The TMO gains can, like in [28], be chosen to be constant, thereby reducing the computational load of the observer. However, a better performance can be achieved by considering the gains as time varying (see [48]). Several options for gain selection are available, while in the following, the Riccati equation will be considered. In general, the method for determining the gain matrices will not affect the observer stability as long as $K$ satisfies the conditions imposed by the TMO stability requirements. The method described here is similar to a Kalman filter, where the gain is chosen based on the solution of the discrete time-varying Riccati equation

$$P_{k|k-1} = \Phi P_{k-1|k-1} \Phi^T + Q$$

$$K_k = P_{k|k-1}(C(P_{k|k-1}C^T + R)^{-1}$$

$$P_{k|k} = (I - K_k C)P_{k|k-1}(I - K_k C)^T + K_k R K_k^T$$

(37)

(38)

(39)

where $k$ is the discrete incrementing index and $\Phi = e^{AT}$ is the discrete transition matrix, with the sample rate and system order denoted as $T$ and $n$, respectively. The dynamics of the TMO is the same as for Kalman filters, and the $P$, $R$, and $Q$ matrices may, therefore, be interpreted as covariance matrices. The error covariance matrix $P$ has elements corresponding to the state vector, and will therefore vary in size when the number of available satellites changes due to the inclusion of the carrier-phase ambiguities. The covariance of the ambiguities $P_N$ is found as the lower right $m - 1 \times m - 1$ submatrix of $P$. The gain selection can be tuned by choosing the measurement $R$ and state covariance matrices $Q$, where it is common to consider them as consisting of the elements pertaining to the covariance of the states or measurements $\sigma^2$ [1, Sec. 14.4.2]

$$Q = \text{diag} (\sigma^2_{p_x}, \sigma^2_{p_y}, \sigma^2_{p_z}, \sigma^2_{r}, Q_{\Delta\Delta})$$

$$R = \text{blockdiag} (R_{\Delta p}, R_{\Delta v}, R_{\Delta b})$$

(40)

(41)

The diagonal elements of the $Q$ matrix correspond to the state vector elements describing the expected level of noise in the state. Similarly, for the $R$ matrix, the diagonal elements correspond to variances of the aiding DD satellite measurements. In [1, Sec. 9.4.2.4], a time-varying $R$-matrix
is proposed, where the coefficients are determined based on the satellite elevation and range acceleration. Here, the matrices (40) and (41) are kept constant. Due to the differencing with a reference satellite, \( R_{\Delta x} \) has correlation elements, such that the covariance matrix is not diagonal. The submatrices of \( R \) are, in general, given as

\[
R_{\Delta x} = \begin{bmatrix}
\sigma_{\Delta x}^2 & \frac{1}{2} \sigma_{\Delta x}^2 & \frac{1}{2} \sigma_{\Delta x}^2 & \cdots & \frac{1}{2} \sigma_{\Delta x}^2 \\
\frac{1}{2} \sigma_{\Delta x}^2 & \sigma_{\Delta x}^2 & \frac{1}{2} \sigma_{\Delta x}^2 & \cdots & \frac{1}{2} \sigma_{\Delta x}^2 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\frac{1}{2} \sigma_{\Delta x}^2 & \frac{1}{2} \sigma_{\Delta x}^2 & \sigma_{\Delta x}^2 & \cdots & \frac{1}{2} \sigma_{\Delta x}^2 \\
\frac{1}{2} \sigma_{\Delta x}^2 & \frac{1}{2} \sigma_{\Delta x}^2 & \frac{1}{2} \sigma_{\Delta x}^2 & \cdots & \sigma_{\Delta x}^2 
\end{bmatrix}
\tag{42}
\]

where \( * \) is a placeholder for pseudorange, carrier-phase, or carrier-phase derived Doppler measurements. \( Q_{\Delta x\Delta N} \) is constructed in a similar way, where the diagonal is \( \sigma_{\Delta x \Delta N} \) with the remaining elements being \( 1/2 \sigma_{\Delta x \Delta N} \).

It is possible to reduce the computational load without jeopardizing the performance significantly by determining the TMO gains on a slower time scale, as shown in [27]. The observer can be implemented using the corrector–predictor architecture of [50, Sec. 11.3.4], where the observer estimates are propagated at IMU frequency and corrected at the lower GNSS receiver frequency whenever satellite measurements are available. The gain selection can then be implemented on a third and slower time scale where the gains are updated for every 100–2,000 GNSS correction. This slower time scale can be chosen, since the \( C \) matrix is slowly time varying.

### B. Integer Ambiguities

The ambiguities in the DD observer are initially considered real valued, and however, the precision can be further improved if they are correctly resolved to integers. Several methods for fixing the carrier-phase ambiguity have been proposed. Here, the “fix and hold” method from [1] will be used to fix the combined variable \( \nabla \Delta \hat{N} := [\nabla \Delta \hat{N}_1; \nabla \Delta \hat{N}_2; \ldots; \nabla \Delta \hat{N}_{m-1}] \) to integer values.

The initial estimate of the ambiguity vector can be determined as the difference between carrier-phase and pseudorange measurements, i.e., subtraction of (12) from (13)

\[
\nabla \Delta \hat{N} = \frac{1}{\hat{\lambda}} (\nabla \Delta \phi - \nabla \Delta \rho) \tag{43}
\]

where \( \nabla \Delta \phi := [\nabla \Delta \phi_1; \nabla \Delta \phi_2; \ldots; \nabla \Delta \phi_{m-1}] \) and \( \nabla \Delta \rho := [\nabla \Delta \rho_1; \nabla \Delta \rho_2; \ldots; \nabla \Delta \rho_{m-1}] \) are vectors of combined variables. The initialization offered by (43) can be used when new satellites are introduced to the constellation. Initialization should also be carried out if a satellite is reintroduced after a period of obstruction or loss-of-lock, as the ambiguity will have changed.

The initial ambiguity estimates depend strongly on a good initial position of the rover and base station. As these might be difficult to obtain prior to flight, the ambiguities should be iterated by the TMO before trying to fix to integer value, to decrease the risk of fixing to wrong integers. After initialization, the ambiguities will, therefore, be propagated by the TMO as real-valued estimates. The estimate can be tested for convergence to integer values by minimizing [1]

\[
\Omega = \min_{\nabla \Delta \hat{N} \in \mathbb{Z}^{m-1}} (\nabla \Delta \hat{N} - \nabla \Delta \hat{N})^T P_N^{-1} (\nabla \Delta \hat{N} - \nabla \Delta \hat{N}) \tag{44}
\]

where \( P_N \in \mathbb{R}^{m-1 \times m-1} \) is the covariance matrix of the ambiguities, and \( \nabla \Delta \hat{N} \) is an integer candidate vector. The candidate vector belongs to the search space of \( S := (\nabla \Delta \hat{N} \in \mathbb{Z}^{m-1}| \nabla \Delta \hat{N} = \nabla \Delta \hat{N} + \sigma_N) \) where \( \sigma_N \) is the variance of the ambiguity estimates determined by the diagonal elements of \( P_N, \sigma_N = \text{diag}(P_N)^{1/2} \). The constant \( \sigma_r \) denotes the confidence interval, which in the following will be considered as \( \sigma_r = 3.29 \) for 99.9% confidence interval, assuming normal distribution.

All possible candidate vectors in the search space are tested to see which minimizes (44). The relation between the smallest and second smallest values of (44), respectively, called \( \Omega_1 \) and \( \Omega_2 \), is used to determine whether significant convergence to a candidate vector is achieved. Here, significant is used in a covariance sense, as the search space depends on the covariance matrix of the ambiguities. The convergence test consists of verifying that the best solution \( \Omega_1 \) is sufficiently far from the next best solution

\[
\Omega_2 / \Omega_1^{1/2} \geq t_N \tag{45}
\]

where \( t_N \) is a threshold value. If the ratio is larger than the threshold, the test is accepted and the ambiguities are fixed to the candidate set corresponding to \( \Omega_1 \). This test can be carried out at every observer iteration until the ambiguities are fixed. However, the search space \( S \) will initially be too large to computationally feasibly search through all candidate sets, which encourages the use of methods, such as the LAMBDA method to decrease the search space. The LAMBDA method was proposed in [9]–[11] and uses a change in variables to transform the confidence interval of the ambiguities to cover a smaller area. The transformation is carried out by use of the decorrelation matrix \( Z \) [1, Appendix G]

\[
\hat{x}_Z = Z \hat{x}_N, \quad P_Z = Z P_N Z^T \tag{46}
\]

where \( \hat{x}_N \) is a collection of the DD ambiguity estimates, i.e., \( \hat{x}_N = \nabla \Delta \hat{N} \). The Cholesky factorization is used to determine the decorrelation matrix \( P_N = LDL^\top \), where \( D \) is a diagonal matrix and \( L \) is a lower triangular matrix with unit elements in the diagonal. The decorrelation matrix is then found as \( Z = [L^{-1}]^\top \), i.e., the rounded inverse of the lower triangular matrix.

After the ambiguity transformation, the search space has decreased in size, leaving fewer candidate sets to be tested, making for a faster evaluation. The LAMBDA method is often used in connection with other resolution methods, such as the geometry-based approach. Once the ambiguities have been fixed, the transformation can be reversed. A key feature of this approach is that due to the rounding operation when determining \( Z \), an integer valued \( x_N \) corresponds to an integer valued \( x_Z \), and vice versa. Implementation aspects of the LAMBDA method are considered in [51] and [52].
There are two alternatives for handling the fixed ambiguities: 1) once the ambiguities are fixed to integer values, they are removed from the state vector, decreasing the $A$, $C$, $K$, $Q$, and $P$ matrices in size or 2) the fixed ambiguities are introduced as an additional measurement augmenting the $R$ matrix and the injection terms. The result will be very similar for the two methods. Method 1) might be preferred to 2) for the decrease in computational power.

According to [53], the residual measurement error should be less than 25% of a wavelength for the integer ambiguity resolution to have high probability of resolving to the correct integer. This implies that with high noise levels in the phase measurements or uncertainty in the estimated geometric distance, the resolution of the ambiguities might lead to the wrong integers resulting in a position error. In the event of loss-of-lock of satellite signals, the code measurements can be smoothed by use of the raw Doppler measurements to reduce the convergence time.

VII. EXPERIMENTAL RESULTS

Experimental data from a flight with a fixed-wing Penguin B UAV (see Fig. 4) is used to verify the proposed DD observer structure. The flight was carried out at Eggemoen airport in Norway ($60^\circ12'52''/N$, $10^\circ19'07''/E$). The UAV was equipped with an ADIS 16488 IMU (with internal magnetometer) collecting acceleration, rotation rate, and magnetic field data at 410 Hz, while a u-Blox LEA-6T receiver gathered navigation satellite measurements at 5 Hz coinciding with measurements obtained at a stationary base station at the airport. The base station receiver is of the same brand and type as the rover receiver. The base station additionally logged the ephemeris data required for satellite position and velocity computation. Throughout the flight, the baseline is kept under 1 km, such that the assumption of the atmospheric delays experienced by the two receivers are spatially correlated holds.

The sensors (IMU, GNSS receiver, and antenna) used are considered low cost and while even cheaper models are available, these are thought to represent the general level of sensors used for UAV flights. The resolution of the sensors and the stability characteristics of the inertial sensors is seen as a limitation on the performance. A better performance can be expected by including better inertial sensors, especially increasing the performance of the navigation solution when GNSS signals are obstructed. Higher grade GNSS receivers might decrease the experienced measurement noise, supply higher $C/N_0$ ratios, and offer higher sample rates. If the measurements are not accurately, synchronized errors in the GNSS/INS integration will arise, as investigated in [54]. Here, synchronization is achieved with custom printed circuit boards and a PIC32 microcontroller, for details on the highly accurate timing platform (see [55]).

In the following, two versions of the presented observer structure, an RTK reference solution is computed using the open source RTKLIB. No cycle slips were presented allowing the reference to maintain resolved ambiguities throughout the flight, thereby ensuring centimeter-level accuracy. The initial 2.3% of the flight has unresolved ambiguities resulting in a reference with decimeter-level accuracy, while the remainder of the flight has centimeter-level accuracy. The RTK solution is determined from the same GNSS data as is available to the observer implementations. However, the sampling rate of the reference solution will be at GNSS receiver sampling rate, while the observer estimates will be at the IMU sampling rate. The observer solution will, therefore, be downsampled when compared.

The tuning parameters for the $Q$ matrix are chosen as: $Q = \text{blockdiag}(0.1[m^2], I_3[m^2/s^2], 0.00025 I_3[m^2/s^4], 0.3[m^2], Q_{\Delta N})$, where $\sigma_{\Delta N} = 0.01[\lambda^2]$, while the parameters for $R$ are chosen based on the STDs of Table I: $\sigma_{\Delta \rho} = 1.1[m^2], \sigma_{\Delta v} = 0.03[m^2/s], \sigma_{\Delta \upsilon} = 0.06[m/s^2]$ increased to accommodate for bounding Gaussian white noise and higher sample rate as discussed in [1, Sec. 9.4.2.4]. The pseudorange STD has been decreased compared with Table I to ensure faster convergence. Units have been given in square parenthesis such as not to confuse with variables or constants. The remaining parameters are $M_0 = 0.0087[rad^2/s^2]$, $k_1 = 0.8$, $k_2 = 0.2$, $k_I = 0.004$, and $\lambda = 0.1903[\lambda]$ corresponding to the GPS L1 wavelength. The ambiguity threshold is chosen as $\tau_N = 3$, as suggested by RTKLIB.

The position of the base station is initialized as an average of the receiver position measurements determined over 25 min of logged data. This is considered to be sufficiently accurate for this application. However, more accurate initial position estimates will increase the transient performance, encouraging to average over longer periods when possible. PPP can ensure a decimeter accurate estimate of the base station, and however, the drawbacks are long initialization time and access to precise ephemeris and atmospheric data. The initialization process presented in [46] can be used to initialize the position of the rover and base station to ensure that the initial positions are close to the true position.
Fig. 5. Sky plot showing elevation and azimuth of satellites. The satellite IDs are included for ease of reference.

In Fig. 5, the sky plot of the constellation throughout the experiment is shown. The satellite trajectories are marked with individual colors and the last position has been marked with the satellite ID indicating the direction. The reference satellite is chosen to be the SV1 (shown in red line) due to the high initial elevation to minimize atmospheric delays. A minimum elevation requirement of 15° is enforced for dismissal of satellites with high propagation errors.

The rover GNSS receiver antenna is placed on top of the UAV between the wings. Due to the high dynamics of the flight with sharp turns, the rover loses lock on satellites with low elevation. Only five satellites are considered in the following test to ensure that the integer ambiguities do not require frequent resetting. The satellites to be included are chosen as SV1, SV4, SV11, SV17, SV20, and SV32, which have high elevation angles. The dilution of precision (DOP) will be large due to the clustering of the used satellites, and higher performance, especially in the vertical component, might be expected when using a broader satellite constellation.

Here, the constellation can be considered akin to a worst case scenario with regard to the DOP.

The trajectory of the UAV includes several fast dynamic maneuvers, such as figures of eights and circles with radius of approximately 210 m. Some points of interest during the flight will be marked in the subsequent figures; ➀ denotes take-off, ➁ notes when the UAV has reached the desired altitude beginning a figure-of-eight flight sequence, at ➂, the flight pattern is changed to circles, followed by ➃ marking the descent with the reference reverting to lower accuracy making it unfit for further comparison shortly prior to landing. These points of interest will be marked in Figs. 6–8, 11, and 12 for ease of performance comparison throughout the flight. The take-off, initial trajectory, and landing were performed with manual control, whereas the figures of eight and circles were operated by the autopilot.

The initial part of the trajectory can be seen in Fig. 6 showing the flight path up to ➁, while the attitude estimation is shown in Fig. 7 with the points of interest marked. The attitude is shown for both cases and is converted from quaternions to Euler angles describing the rotation from Body- to NED-frame for more intuitive understanding. The attitude estimation is similar for both cases considered, which is attributed to the modular observer structure where only the specific force estimate changes with the case. The point of Fig. 7 is not to be able to compare the attitude estimation between the two cases but rather to visualize that there is a little difference between the cases. It should be noted that the oscillatory behavior of the pitch estimate between ➂ and ➃ corresponds well with the observed behavior during flight, where a slightly sinusoidal vertical trajectory during the circular flight pattern was present. This was due to an uncompensated issue in the autopilot.

The relative position between rover and base station is determined and compared with the relative position offered by the reference. The relative position estimation errors for the proposed observer can be seen in Fig. 8. It is clear that the integer ambiguities are resolved well in advance of take-off. The transient period of 35 s is not visible in Fig. 8, due to the initial offset. However, the transient behavior can be seen in Fig. 9. The performance of the position estimates is compared in Table II, where root-mean-square errors (RMSEs)
and STDs are summarized. The values are determined after the initial transient period of 35 s. An additional test case is included in Table II, where the base station position has been excluded from the state vector and is considered constant at the initial position. This test will be denoted Case C and will utilize pseudorange and carrier-phase aiding. The accuracy obtained by Case C is greatly dependent on the quality of the base station position estimate, while Cases A and B estimate the base station position as part of the state vector.

The performance of Cases A and Case B is very similar and is not influenced by the dynamics of the flight. Case B might have advantages in velocity estimation, and however, as no reliable velocity reference is available, this cannot be confirmed. Small drifts are present in the position estimation, e.g., of approximately $1.7 \times 10^{-5}$ m/s in the North direction.

The fast convergence and high accuracy shown in Fig. 8 are attributed to the ambiguity resolution. In Fig. 10, the ambiguities for Case A are shown for the first 35 s, with similar results obtained for Case B. The real-valued estimate is shown in blue line, with the RTKLIB reference integers in black line. The rounded estimates are shown in red line and only serve as a visual comparison to the reference integers, as the resolved integers are not guaranteed to be the same as the rounded real-valued estimates. The estimated ambiguities are seen to quickly converge to the reference value, with one exception for SV17, where the stationary value is one wavelength off. This is considered acceptable and will still give high precision, as seen in Fig. 8 and Table II.

Looking at Table II, the RMSE and STDs are seen to be on centimeter level for all three cases, with subcentimeter-level STD for the horizontal components. The vertical components are seen to be less accurate than the horizontal components for all three cases. This can be attributed in part to the clustered satellite constellation, and the general result of less accurate vertical channel offered by GNSS measurements. Case C has slightly a better performance than Case A. However, as the difference is on submillimeter level, Cases A and C practically offer the same solution. It is, therefore, concluded that the inclusion of the base station position in the state vector does not lead to poorer state estimates and will be encouraged, especially in case of a slightly moving base station such as a base station on a ship during station keeping, or if a precise initial base station position cannot be obtained. These performance results are considered good in view of the low-cost sensors used and the clustered satellite constellation.

### Table II

<table>
<thead>
<tr>
<th>AIDING</th>
<th>RMSE</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>$\varphi$</td>
<td>$\nu$</td>
</tr>
<tr>
<td>Case A:</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Case B:</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Case C:</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
A. Float Versus Fixed Ambiguities

In order to ascertain the influence of the ambiguities, a test with only real-valued ambiguities is compared with Case A. This test will be denoted Case D, and will not differ in any way from Case A, apart from the carrier-phase ambiguities being unresolved. In Fig. 11, Cases D and A are compared. For Case D, the RMSE value is $[1.277; 0.942; 4.399]$ m with the STDs of $[1.212; 0.865; 1.711]$ m, which is approximately two orders of magnitude larger than with resolved ambiguities. It is clear that for applications with demands of high accuracy, the ambiguities need to be resolved, which is expected.

B. Low Elevation Constellation

A low elevation test is carried out to evaluate the performance when using satellites with low elevation angles. This can simulate a high latitude flight close to the poles where the GPS constellation coverage is poor granting only access to satellites at low elevation. The same data as used in the above test are utilized, here with an elevation mask of $5^\circ$. The atmospheric disturbances experienced will differ compared with a high latitude flight, and however, this test can give an impression on the possible performance achievable.

In a high latitude, e.g., Arctic, setting the environment would be different from the one considered here where major differences include: 1) the terrain in the Arctic will be completely open sky, whereas here some obstruction is caused by trees and mountains; 2) the ionospheric and tropospheric conditions; 3) a different number of satellites; and 4) choice of satellites. The best set of satellites based on, e.g., DOP and $C/N_0$ will be selected, whereas here a poor subset of the satellites is chosen.

Satellites with high elevation are masked out. The satellites, such as SV4, SV6, SV11, SV14, SV23, and SV31, are used, where SV11 has been included to satisfy the $m \geq 4$ assumption, as the low elevation satellites often are obstructed jeopardizing the constellation size assumption. The satellite SV12 is excluded as several cycle slips were detected.

The ambiguity estimates are reinitialized if a satellite is reintroduced in the constellation.

The proposed observer without carrier-phase derived Doppler aiding (as in Case A) is tested with the low elevation constellation, where the performance is shown in Fig. 12. Instances when $m < 5$ are marked with vertical gray lines, e.g., at time 1085 s. The test is shortened by an approximately 45 s, compared with the previous tests, due to extensive loss of low elevation satellites toward the end of the flight.

It is not possible to resolve the ambiguities reliably due to frequent loss of satellite lock and reintroduction of the satellites. This has a clear impact on the performance of the position estimation. The RMSE values are $[2.200; 6.001; 3.961]$ with STDs $[0.367; 0.821; 1.909]$, which is significantly larger than the values listed in Table II and for the real-valued test in Section VII-A. However, the horizontal accuracy for the rover is seen to be on meter level. This is an improvement over standalone solutions, where larger deviations would be expected for a low elevation constellation.

VIII. Conclusion

A tightly coupled GNSS/INS integration scheme using a dual GNSS-receiver configuration between a stationary base station and a moving rover was proposed, where the position and linear velocity of a rover were estimated by a proposed TMO, while attitude was determined with a recent nonlinear attitude observer. The proposed TMO utilizes the error of the DD pseudorange, carrier-phase, and carrier-phase derived Doppler satellite measurements between the receivers in the injection terms.

The observer was verified using flight data from a UAV equipped with low-cost sensors, where position estimates were shown to be within 2–4 cm of the RTK reference solution throughout a 20-min-long flight. Convergence of the ambiguities introduced by the carrier-phase measurements was achieved by inclusion in the state vector, where they were initially considered real valued later to be resolved to integer values.
A. Proof: TMO Injection Terms

The DD injection terms are expressed as
\[
e_{γ,ij} = \| p_{γ} - p_{γ}^t \|_2 - \| p_{γ} - p_{γ}^t \|_2 + \| \tilde{p}_{γ} - p_{γ}^t \|_2 - \| \tilde{p}_{γ} - p_{γ}^t \|_2
\]
by substitution of \( r \) with \( s \). The higher order terms can be bounded, as shown in [46]
\[
\| \eta_{p,ij} \|_2 \leq \eta_{r} \| \tilde{p}_{r} \|_2 + \eta_{s} \| \tilde{p}_{s} \|_2
\]
where
\[
\eta_{r} = \frac{\| \eta_{r,i,j} \|}{\| \psi_{r} \|}, \quad \eta_{s} = \frac{\| \eta_{s,i,j} \|}{\| \psi_{s} \|}
\]
and
\[
\psi_{r} = \| \psi_{r,i,j} \|, \quad \psi_{s} = \| \psi_{s,i,j} \|
\]
with \( \psi_{r,i,j} \) and \( \psi_{s,i,j} \) the lower bounds on the geometric distance between receiver and \( i \)th satellite, and \( \| v_{e} - v_{e}^t \|_2 \leq \psi_{r} \).

ACKNOWLEDGMENT

The authors would like to thank the UAV engineers at the Norwegian University of Science and Technology (NTNU) and Maritime Robotics AS for their assistance, in particular L. Semb and C. E. Stephansen. They would also like to thank the team at NTNU for the significant contributions to the construction of the UAV payload, in particular S. Albrektsen.

REFERENCES


Tor Arne Johansen received the M.Sc. and Ph.D. degrees from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway.
He was a Researcher with SINTEF Electronics, Trondheim, and Cybernetics. He was appointed as an Associate Professor with NTNU in 1997, and promoted to a Professor in 2001. In 2002, he co-founded the company Marine Cybernetics AS, Trondheim, where he was a Vice President until 2008. He is currently a Principal Researcher with the Center of Excellence on Autonomous Marine Operations and Systems and the Director of the Unmanned Aerial Vehicles Laboratory, Department of Engineering Cybernetics, NTNU. He has published more than 100 articles in international journals and numerous conference articles and book chapters in the areas of control, estimation, and optimization with applications in the marine, automotive, biomedical, and process industries.

Nadezda Sokolova received the M.Sc. degree in geomatics engineering from the University of Calgary, AB, Canada, the M.Sc. degree in space and aeronautical engineering from Narvik University, Narvik, Norway, and the Ph.D. degree from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, with a focus on weak global navigation satellite system (GNSS) signal tracking and use of GNSS for precise velocity and acceleration determination.
She is currently a Research Scientist with SINTEF, Trondheim, and an Adjunct Associate Professor with the Engineering Cybernetics Department, NTNU, where she is involved in multifrequency, multiconstellation ground-based augmentation system design for precision landing operations, and multisensor navigation.

Thor I. Fossen (F’16) received the M.Sc. degree in marine technology and the Ph.D. degree in engineering cybernetics from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 1987 and 1991, respectively.
From 1989 to 1990, he was a Fulbright Scholar in aerodynamics and flight control with the Department of Aeronautics and Astronautics, University of Washington, Seattle, WA, USA. In 1993, he was appointed as a Professor of guidance, navigation and control at NTNU. He is currently the Co-Director of the CoE Centre for Autonomous Marine Operations and Systems, NTNU. He has authored approximately 400 scientific papers and six textbooks.
Dr. Fossen received the Automatica Prize Paper Award for the research on the weather optimal positioning control of marine vessels in 2002. He has served as an Associate Editor for Automatica. He is currently an Associate Editor of the Control Engineering Practice.