
Mathieu Joerger, Jason Neale and Boris Pervan
Illinois Institute of Technology
Chicago, Illinois, US
joermat@iit.edu

Seebany Datta-Barua
Atmospheric and Space Technology Research Associates (ASTRA), LLC
San Antonio, Texas, US

Abstract—The integration of ranging signals from multiple satellite constellations opens the possibility for rapid, robust and accurate positioning over wide areas. Algorithms for the simultaneous estimation of carrier phase cycle ambiguities and user position and for the detection of faults over a fixed smoothing time-interval were derived in previous work. For high-integrity precision applications, ensuring the robustness of measurement error and fault-models is an exacting task, especially when considering sequences of observations. In this research, a new RAIM-based approach is established, which aims at directly determining the worst-case single-satellite fault profile. Also, the robustness of newly derived ionospheric error models is experimentally evaluated using dual-frequency GPS data collected over several months at multiple locations. An integrity analysis is devised to quantify the impact of traveling ionospheric disturbances (TIDs) on the final user position solution. Finally, overall navigation system performance is assessed for various combinations of GPS, Galileo and low earth orbiting Iridium satellite signals.

Keywords: robust modeling; fault detection; multi-constellation; Iridium; traveling ionospheric disturbance (TID)

I. INTRODUCTION

This paper describes the design, analysis and evaluation of new fault-detection algorithms and measurement error models for high-integrity multi-constellation navigation systems. The proposed algorithms exploit multi-constellation signal redundancy and geometric diversity to perform rapid floating carrier phase cycle ambiguity estimation, and fault-detection using carrier phase Receiver Autonomous Integrity Monitoring (RAIM). In addition, robust measurement error models are established to enable precise non-differential navigation using low-broadcast-rate wide-area corrections. An overall system integrity analysis is devised and implemented. It explores the potential of multi-constellation navigation systems to provide robust carrier phase positioning over continental areas.

Carrier phase positioning is contingent upon estimation (or resolution) of cycle ambiguities. These remain constant as long as they are continuously tracked by the receiver. An efficient solution for their estimation is to use the bias observability provided by redundant satellite motion [1]. This principle was demonstrated to be particularly effective in navigation systems that augment GPS with ranging signals from fast moving low earth orbit (LEO) spacecraft [2].

Estimation and detection algorithms were designed in [3] to fully exploit angular variations in satellite lines of sight (LOS). A fixed-interval smoothing filter was established for combined estimation of user position and floating cycle ambiguities. Fault-detection was achieved using a batch least-squares residual-based carrier phase RAIM method. The smoothing interval \( T_s \) was limited to about 10 min because most measurement errors can be reliably modeled over such short durations. The algorithms were evaluated with sequences of code and carrier phase measurements from GPS and from LEO Iridium telecommunication space vehicles (SV).

However, one major challenge in validating these algorithms for high-integrity precision applications is ensuring the robustness of measurement error and fault models. Time-processing (needed for cycle ambiguity estimation) introduces yet another dimension to the problem in that the dynamics over time of the errors and faults must be carefully accounted for. In response, in this work, new RAIM-based procedures are derived, and measurement error models are refined and tested against large sets of experimental data.

First, a method is established to compute the worst-case fault detection performance. All measurements collected during the smoothing period are vulnerable to rare-event integrity threats such as satellite or equipment failures. In order to capture the impact of such failures over time, a set of realistic fault modes can be considered (such as impulses, steps, and ramps of all magnitudes and start times). But this set of canonical faults does not constitute a comprehensive description of all integrity threats. Therefore, a new concept is introduced with the derivation of theoretical worst-case faults, which maximize the probability of missed detection. In parallel, careful analysis of missed-detection mechanisms identifies cases of poor satellite geometries and motivates the derivation of a new “extended-window” RAIM-based algorithm. The extended-window method enhances the current-time fault-detection capability by exploiting stored observations from improved past-time geometries. It is particularly efficient with fast changing LEO satellite geometries.

The authors gratefully acknowledge The Boeing Co. and the Naval Research Laboratory for sponsoring this work. However, the opinions expressed in this paper do not necessarily represent those of any other organization or person.
The second part of this paper revisits measurement error models derived in [4]. A nominal configuration for a notional near-future LEO satellite-augmented GPS navigation system was assumed, for single-frequency users receiving low-rate corrections from a network of wide-area ground stations (e.g., similar to the Wide Area Augmentation System or WAAS). The resulting measurement equation expressed measurement dependency on error sources including satellite clock and orbit ephemeris, ionospheric and tropospheric refraction, multipath and receiver noise.

For non-differential single-frequency implementations, the ionosphere is the most influential source of error. In this research, an ionospheric error analysis is performed using experimental data from Continuously Operating Reference Stations (CORS) collected over several months at multiple locations. The analysis shows that a piecewise linear model of the vertical ionospheric delay captures most of the ionospheric variations at mid-latitudes. However, the ionosphere also exhibits localized wave-like structures causing decimeter-level errors on low-elevation SV signals. These structures are often referred to as Traveling Ionospheric Disturbances (TIDs) [5]. With the assumed nominal system configuration, TIDs will significantly alter GPS and LEO satellite ranging observations. In this research, two strategies are employed to mitigate the effect of TIDs. One is to derive a new measurement error model with a sinusoidal term that accounts for TIDs. The other is a conservative procedure to include mis-modeling errors with unknown time-behavior in the estimation and detection algorithms. Both approaches are calibrated using experimental data.

Performance evaluations are structured around benchmark requirements inspired from civilian aviation standards. For example, a predefined 10 m vertical alert limit (VAL) is considered for the purpose of the analysis. An example application of aircraft precision approach is implemented. Fault-free (FF) integrity is evaluated by covariance analysis, and RAIM detection performance is quantified for a set of realistic single-satellite faults (SSF) and for theoretical worst-case SSF. The fault-detection capability of the extended-window algorithm is compared to the original RAIM method. A sensitivity analysis of the combined FF and SSF performance assesses the influence of new parameters in the updated ionospheric error model. In particular, it reveals the decisive impact of TIDs on user positioning performance.

Section II of this paper describes the estimation and detection algorithms: the extended-window RAIM method is introduced, worst-case faults are derived, and a conservative approach for incorporating mis-modeling errors is presented. In Sec. III and IV, ionospheric error models are defined and evaluated using experimental data from CORS. Also, a refined model of the satellite orbit ephemeris error is given in Sec. V. It accounts for non-radial components of the orbit error, which are negligible for GPS medium earth orbit (MEO) SVs but not for LEO spacecraft. The framework for the performance analysis is described in Sec. VI, including details of the navigation system architecture, the complete measurement equation and the benchmark mission and requirement definition. Finally, the overall system performance is established in Sec. VII for near future single-frequency implementations and for longer-term future, dual-frequency, multi-constellation navigation systems including combined GPS/Galileo, Iridium/GPS and Iridium/GPS/Galileo constellations.

II. ESTIMATION AND DETECTION ALGORITHMS

A. Fixed-Interval Estimation Algorithm

A fixed-interval smoothing algorithm was devised in [3] for the simultaneous estimation of user position and of floating carrier phase cycle ambiguities. It is compatible with real-time implementations provided that sufficient memory is allocated to the storage of a finite number of past measurements and LOS coefficients (collected over the smoothing period \( T_F \)).

Current-time (and past-time) optimal state estimates are obtained from iteratively feeding the stored finite sequence of observations into a forward-backward smoother. Past-time estimates are needed later in the RAIM-based procedure for residual generation (next subsection). The smoother is equivalent to a batch measurement processing method, which is presented below for clarity in exposition.

Consider first the vector of carrier phase observations for a satellite \( s \) in view between epochs 0 and \( n_p - 1 \) (0 and \( n_p - 1 \) are the first and last epochs of the smoothing interval if satellite \( s \) is visible during the entire interval \( T_F \)):

\[
\Phi = \begin{bmatrix}
\phi_0 & \cdots & \phi_{n_p - 1}
\end{bmatrix}^T.
\]  

(1)

Carrier phase observations for all \( n_s \) visible satellites are then stacked together:

\[
\Phi = \begin{bmatrix}
\phi_1^T & \cdots & \phi_n^T
\end{bmatrix}^T.
\]  

(2)

The complete sequence of code \( p \) and carrier phase \( \phi \) signals for all satellites over the smoothing interval \( T_F \) are included into a batch measurement vector:

\[
Z = \begin{bmatrix}
\Phi^T & \rho^T
\end{bmatrix}^T,
\]  

(3)

which can be expressed as a state space realization:

\[
Z = Hx + v,
\]  

(4)

where the state vector \( x \) is such that:

\[
x = \begin{bmatrix}
u_0^T & \cdots & u_{n_p - 1}^T & N^T & x_{EFF}^T
\end{bmatrix}^T.
\]  

(5)

The dynamics of the user position and receiver clock deviation vector \( u_k \) are unknown. Different states are therefore allocated to the vector \( u_k \) at each time step \( k \), as opposed to the other parameters that are modeled as constants over interval \( T_F \). The vector \( N \) is comprised of cycle ambiguity states for
State augmentation is used to incorporate the dynamics of the measurement error models. Thus, the error state vector $x_{\text{ERR}}$ is made of the constant parameters of the error models. A detailed description of $x_{\text{ERR}}$ is given in Sec. VI and is not needed for the next steps of the algorithm derivation. Prior knowledge on the error state vector $x_{\text{ERR}}$ is expressed in terms of bounding values on each element’s probability distributions, and included in the system with the a-priori information matrix $P_{\text{PK}}^{-1}$. Finally, the measurement noise vector $v$ in (4), with covariance matrix $V$, is utilized to introduce the time-correlated noise due to multipath as well as receiver noise. Detailed derivations of the measurement covariance and observation matrices (respectively $V$ and $H$) are given in Appendix.

The weighted least squares state estimate $\hat{x}$ is obtained using the equation (e.g., [6]):

$$\hat{x} = S \ z,$$

(7)

where $S$ is the weighted pseudo-inverse of $H$:

$$S = P_{\text{PK}}^{-1} H^T V^{-1}.$$

(8)

The weighted least-squares state covariance matrix $P_{\text{x}}$, which includes prior knowledge on a subset of states, is expressed as:

$$P_{\text{x}} = \left( \begin{bmatrix} 0 & 0 \\ 0 & P_{\text{PK}}^{-1} \end{bmatrix} + H^T V^{-1} H \right)^{-1}.$$

(9)

If the focus of the estimation performance analysis is on a subset of states, for example on the current-time vertical position state $x_U$, a transformation matrix $T_U$ can be defined as:

$$T_U = \begin{bmatrix} \mathbf{0}_{n_{\text{x}} \times 1} & \mathbf{1} \end{bmatrix},$$

(10)

where, following the order in which the $n_x$ states are arranged in $x$, $n_{\text{x}}$ and $n_{\text{y}}$ are the numbers of states respectively before and after $x_U$ (so that $n_{\text{x}} = n_{\text{x}} + 1 + n_{\text{y}}$). The diagonal element of $P_{\text{x}}$ corresponding to the current-time vertical position covariance is noted $\sigma_{U}^2$ and can be expressed as:

$$\sigma_{U}^2 = T_U P_{\text{x}} T_U^T.$$

(11)

The standard deviation $\sigma_U$ is used in Sec. VI to determine FF availability.

### B. Batch Residual RAIM Detection Algorithm

State estimation is based on a history of observations, all of which are vulnerable to satellite faults. To protect the system against abnormal events, a RAIM-type process is implemented, using the least-squares residuals of the batch measurement equation (4). The least-squares residual RAIM methodology [7] gives a statistical description of the impact of a measurement fault vector $f$ (of same dimensions as $z$), whose non-zero elements introduce deviations from normal FF conditions. Equation (4) becomes:

$$z = H x + v + f,$$

and the batch least-squares residual vector is defined as:

$$r \equiv (I - H S) z = (I - H S) (v + f),$$

(13)

where $I$ is the identity matrix of appropriate size.

The residual-based RAIM methodology investigates the impact of the fault vector $f$:

- on the current-time vertical position state estimate error $\delta x_U$, given by:

$$\delta x_U = x_U - \hat{x}_U = T_U S (v + f)$$

(14)

- and on the norm $\|r\|_{\text{wr}}^2$ of $r$ weighted by the measurement noise information matrix $V^{-1}$, which is defined as:

$$\|r\|_{\text{wr}}^2 = r^T V^{-1} r.$$  

(15)

The estimate error $\delta x_U$ is normally distributed with mean $\delta x_U$ and variance $\sigma_U^2$. The following notation is used in the rest of the paper:

$$\delta x_U \sim N(T_U S f, \sigma_U^2).$$

(16)

The weighted norm $\|r\|_{\text{wr}}^2$ of $r$ follows a non-central chi-square distribution with $n_x - (4n_p + n_s)$ degrees of freedom (where $n_x$, $n_p$, and $n_s$ are the numbers of measurements, of epochs in the batch and of visible satellites) and non-centrality parameter $\lambda_{\text{SSF}}^2$ given by [8]:

$$\lambda_{\text{SSF}}^2 = f^T V^{-1} (I - H S) f$$

(17)

The system is said to produce hazardous information if:

$$|\delta x_U| > \text{VAL}.$$  

(18)

In addition, a detection threshold $R_c$ is set in compliance with a continuity requirement $P_c$ (see Sec. VI for details) to limit
the probability of false alarms under fault free conditions \[9\] 
\[ R_C^2 \] is the value for which the chi-square cumulative distribution function with \( n_Z - (4n_s + n_x) \) degrees of freedom equals 1 – \( P_c \). As a result, a measurement failure is undetected if:

\[ \|r\|_w < R_C. \]  \hspace{1cm} (19)

The probability of missed detection \( P_{MD} \) is defined as a joint probability:

\[ P_{MD} = P\left( \|\delta x_c\| > VAL \ , \ \|r\|_w < R_C \right). \]  \hspace{1cm} (20)

Finally, \( P_{MD} \) is used in Sec. VI to determine SSF availability.

C. Extended Window RAIM Algorithm

If the fault-detection requirement is not met using the RAIM method and if the mission duration exceeds the smoothing interval \( T_F \), the extended window RAIM algorithm may be implemented. This procedure was inspired from a previous Relative RAIM (RRAIM) method derived in [10]. The concept of RRAIM was first introduced in the context of the Global Navigation Satellite System (GNSS) evolutionary architecture study (GEAS) [11]. RRAIM assumes that measurement integrity is ensured at some prior time \( t_0 \) (either by the ground segment, or by a previous RAIM verification). The precise and unambiguous relative change in carrier phase observations between the current epoch \( t_j \) and a reference epoch \( t_0 \) is exploited for measurement integrity monitoring at time \( t_j \). In other words, RRAIM enables to coast through temporary periods of poor satellite geometries. It was demonstrated to be particularly efficient with the fast changing Iridium/GPS constellation in [10].

However, expressions of the position-domain RRAIM equations are complicated by terms accounting for the change in geometry between \( t_0 \) and \( t_j \). In addition, RRAIM is not optimal for three reasons: (a) the sub-allocation of the continuity requirement \( P_c \) between the initial RAIM check at \( t_0 \) and the RRAIM check at \( t_j \), (b) the overly-conservative choice of a worst tolerable fault vector at epoch \( t_0 \), and (c) the actual fault occurring between \( t_0 \) and \( t_j \) being unconstrained with respect to the tolerable fault at \( t_0 \).

In response, a new approach is proposed with the extended-window RAIM method. It is simpler to implement than RRAIM, but it is computationally more intensive. The extended-window concept stems from two major observations drawn from careful analysis of faults affecting Iridium/GPS signals that could not be detected using the original RAIM process.

- Periods of consecutive unavailable geometries do not exceed two-to-five minutes, i.e., there is always an Iridium/GPS geometry that has been RAIM-validated within the past ten minutes.
- Faults that are the most difficult to detect affect Iridium satellite signals. In fact, most undetected faults are unobservable using the residual because their profile is a scaled version of the Iridium satellite range variation.

The extended window RAIM method aims at coasting from a past-time SSF-available geometry through periods of poor Iridium SV trajectories using GPS measurements.

The extended window RAIM procedure can be thought of as a fault detection and exclusion mechanism, where measurements to exclude (Iridium signals) have been identified a priori. First, the probability \( P_{MD} \) is computed using the current-time batch of measurements collected between \( t_j - T_F \) and \( t_j \) (nominal RAIM). As illustrated in Fig. 1, if \( P_{MD} \) fails to meet the SSF integrity requirement (defined in Sec. VI), then the batch is extended backwards using stored past measurements. Past Iridium and GPS measurements are considered at previous epochs between \( t_0 - T_F \) and \( t_0 \), but only GPS observations are used between \( t_0 \) and the current epoch \( t_j \). This means that Iridium measurements between \( t_0 \) and \( t_j \) are excluded, hence avoiding the SV trajectory causing missed-detection (dashed area in Fig. 1). The reference epoch \( t_0 \) is selected going backwards in time until the SSF availability criterion (in Sec. VI) can be satisfied.

Figure 1. Time-Line of Signals Used in RAIM and in the Extended Window RAIM Method
The period $T_{EW}$, defined as $t_j - t_0 + T_F$, is limited by the period of validity of the GPS measurement error models. $T_{EW}$ is assumed to be limited to 10-20 min, which will have to be validated experimentally. Derivation of the probability of missed detection $P_{MD}$ in (20) is unchanged (with increased number of measurements, of exposure period to SV faults, and possibly of visible satellites). The analysis of Sec. VII verifies that in cases of poor geometry, the exclusion of few Iridium measurements enhances the fault-detection performance (and negligibly reduces the estimation performance). Overall, the extended window RAIM method improves the current-time fault-detection capability by exploiting more favorable past-time geometries.

D. Worst Case Fault Derivation

Canonical step and ramp-type fault models of all magnitudes and start times were implemented in [4]. They account for the largest part of realistic satellite-related integrity threats, but do not provide a comprehensive description of all potential fault modes. A more direct approach is investigated here by deriving theoretical single-satellite faults specifically designed maximize the probability of missed detection $P_{MD}$.

The worst-case fault maximizes position estimate error (most hazardous) while minimizing the residual (most misleading). Fault vectors that belong to the range space of $\mathbf{H}$ (e.g., $\mathbf{f} = \mathbf{Hx}_w$, for any vector $\mathbf{x}_w$ of same length as $\mathbf{x}$) are strictly undetectable using the residual ($\lambda_{SSF}^2$ in (17) is zero). In fact, the impact of the vector $\mathbf{x}_w$ is entirely transferred onto the state estimate error vector $\delta \mathbf{x}$ (defined in (14) for $\mathbf{x}_w$). In general, it is known that the estimate error $\delta \mathbf{x}$ and the residual $\mathbf{r}$ are derived from orthogonal components of the fault vector $\mathbf{f}$ (see for example, [12]).

A single-satellite fault (or a fault on a small subset of satellites) only affects few elements of the fault vector $\mathbf{f}$. Let $n_{NZ}$ be the number of non-zero elements in $\mathbf{f}$ (i.e., the number of faulty samples). Following the order in which individual satellite observations are stacked in $\mathbf{z}$, let $n_c$ and $n_d$ be the numbers of measurements respectively before and after the non-zero elements (the total number of measurements $n_z$ equals $n_c + n_{NZ} + n_D$). The vector $\mathbf{f}$ may be expressed as:

$$\mathbf{f} = \mathbf{T}_z \mathbf{f}_{NZ}.$$  \hfill (21)

$\mathbf{T}_z$ is a $n_Z \times n_{NZ}$ matrix defined as:

$$\mathbf{T}_z = [\mathbf{0}_{n_z \times n_c}, \mathbf{I}_{n_{NZ}}, \mathbf{0}_{n_z \times n_D}]^T,$$  \hfill (22)

where $\mathbf{I}_n$ is a $n \times n$ identity matrix, and $\mathbf{f}_{NZ}$ is a $n_{NZ} \times 1$ vector of non-zero elements.

Equations (16) and (17) indicate that the fault vector $\mathbf{f}$ affects the mean of $\delta \mathbf{x}_U$ and the non-centrality parameter $\lambda_{SSF}^2$ of the residual $\mathbf{r}$. Their ratio is named the failure mode slope $g_{FM}$, and is expressed as:

$$g_{FM}^2 = \frac{\mathbf{f}_{NZ}^T \mathbf{T}_z^S \mathbf{T}_z^T \mathbf{U}_S \mathbf{S} \mathbf{T}_z \mathbf{f}_{NZ}}{\mathbf{f}_{NZ}^T \mathbf{T}_Z \mathbf{V}^{-1} (1 - \mathbf{H} \mathbf{S}) \mathbf{T}_Z \mathbf{f}_{NZ}},$$  \hfill (23)

In order to determine the direction of vector $\mathbf{f}_{NZ}$ that maximizes $g_{FM}$, a change of variable is performed by defining $\mathbf{f}_{NZ}$ as:

$$\mathbf{f}_{NZ} = (\mathbf{T}_Z^T \mathbf{V}^{-1} (1 - \mathbf{H} \mathbf{S}) \mathbf{T}_Z)^{1/2} \mathbf{f}_{NZ^*}.$$  \hfill (24)

The following notation is used in the next steps of the derivation:

$$\mathbf{M}_Z = (\mathbf{T}_Z^T \mathbf{V}^{-1} (1 - \mathbf{H} \mathbf{S}) \mathbf{T}_Z)^{1/2}.$$  \hfill (25)

The matrix $\mathbf{V}^{-1} (1 - \mathbf{H} \mathbf{S})$ is of rank $n_z - n_x$. The matrix $\mathbf{M}_Z^{-1}$ is $n_{NZ} \times n_{NZ}$ and is full rank for any $\mathbf{T}_Z$ corresponding to a SSF. In this case, $\mathbf{f}_{NZ}$ is given by:

$$\mathbf{f}_{NZ} = \mathbf{M}_Z \mathbf{f}_{NZ^*},$$  \hfill (26)

and the failure mode slope can be rewritten as:

$$g_{FM}^2 = \frac{\mathbf{f}_{NZ^*}^T \mathbf{M}_Z^T \mathbf{M}_Z \mathbf{M}_Z \mathbf{f}_{NZ^*}}{\mathbf{f}_{NZ^*}^T \mathbf{f}_{NZ^*}}.$$  \hfill (27)

where

$$\mathbf{M}_U = \mathbf{T}_U \mathbf{S} \mathbf{T}_Z.$$  \hfill (28)

The vector $\mathbf{f}_{NZ^*}$ that maximizes $g_{FM}^2$ is the eigenvector $\mathbf{v}_{MAX}$ corresponding to the largest eigenvalue of the symmetric matrix $\mathbf{M}_Z^T \mathbf{M}_Z \mathbf{M}_U \mathbf{M}_Z$. Finally, the worst-case fault $\mathbf{f}_{WORST}$ that maximizes the probability of hazardous misleading information is:

$$\mathbf{f}_{WORST} = \mathbf{T}_Z \mathbf{M}_Z \mathbf{v}_{MAX}.$$  \hfill (29)

A similar derivation can be found in [13] for single-epoch faults simultaneously affecting multiple satellites.

E. Conservative Approach to Account for Mis-Modeling Errors

Measurement error models must account for the instantaneous uncertainty at smoother initiation (absolute measurement error) as well as their variations over the smoothing duration (relative error with respect to initialization). Data from CORS are processed in Sec. IV to quantify the residual ionospheric error after differencing the model from the experimental data. Maximum observed values of the residual errors can be measured, but their time-behavior is extremely difficult to model. In this subsection, a
conservative approach is taken to evaluate the impact of mis-modeling errors on estimation and detection performance.

A mis-modeling error vector $\vepsilon$ is added to the FF measurement equation (4):

$$ z = Hx + v + \vepsilon $$ (30)

In the case where mis-modeling errors affect a subset of observations, a transformation matrix $T_{zs}$ can be employed (next paragraph) to isolate non-zero elements of $\vepsilon$. For example, for single-satellite errors affecting the $n_{zS}$ observations of a satellite $s$, the $n_z \times n_{zS}$ matrix $T_{zs}$ is defined as:

$$ T_{zs} = [0_{n_z \times n_s}, I_{n_z}, 0_{n_z \times n_y}]^T. $$ (31)

where, following the order in which measurements are stacked in $z$, $n_{zA}$ and $n_{zB}$ are the numbers of measurements respectively before and after measurements from satellite $s$ ($n_z = n_{zA} + n_{zB} + n_y$). To be conservative in the performance analysis, the satellite whose observations have the largest impact on the states of interest (e.g., on the current-time vertical position coordinate) can be selected using the following criterion:

$$ \max \left( T_{zs}^T S^T T_U^T T_{zs} \right). $$ (32)

Experimental CORS data analysis (Sec. IV) indicates that TIDs cause decimeter-level magnitude error variations with a wide variety of profiles. TIDs are localized, temporary phenomena that rarely affect more than one satellite at a time. Therefore, the impact of TIDs is modeled on a single satellite. Let $\vepsilon_{TID}$ be the $n_{zS} \times 1$ vector of residual errors due to TIDs on measurements from satellite $s$. The FF measurement equation (20) becomes

$$ z = Hx + v + T_{zs} \vepsilon_{TID}. $$ (33)

An upper limit $b$ is established for the impact of $\vepsilon_{TID}$ on the states of interest (e.g., on the current-time vertical position estimate) using Hölder’s inequality [14]:

$$ \left\| T_{zs} S T_{zs} \vepsilon_{TID} \right\| \leq \left\| T_{zs} S T_{zs} \right\| \left\| \vepsilon_{TID} \right\|_\infty $$ (34)

where $\left\| \vepsilon_{TID} \right\|_\infty = \max \left( \left| \vepsilon_{TID,1} \right|, \left| \vepsilon_{TID,2} \right|, \ldots, \left| \vepsilon_{TID,n_{zS}} \right| \right),$

$$ \left\| T_{zs} S T_{zs} \right\| = \sum_{i=1}^{n_{zS}} \left| \left( T_{zs} S T_{zs} \right)_{ii} \right|, $$ (36)

and $(T_{zs} S T_{zs})_{ii}$ is the $i$th element of the $1 \times n_{zS}$ row vector $(T_{zs} S T_{zs})$. The maximum value that any element of $\vepsilon_{TID}$ can take is noted:

$$ \left\| \vepsilon_{TID} \right\|_\infty = b. $$ (37)

The upper bound $b$ is expressed as:

$$ b = \beta_{TID} \sum_{i=1}^{n_{zS}} \left| (T_{zs} S T_{zs})_{ii} \right|. $$ (38)

It is used in Sec. VI to establish a conservative FF performance criterion in the presence of mis-modeling errors.

In addition, for the SSF performance analysis, the mean of the estimate error $\delta x_u$ in (16) is modified to:

$$ \delta x_u \sim N \left( T_u S f + b, \sigma_u^2 \right). $$ (39)

The vector $\vepsilon_{TID}$ in (33) also causes the detection threshold $R_c$ to become non-centrally chi-square distributed. The fault-free non-centrality parameter $\lambda^2_{FF}$ is defined as:

$$ \lambda^2_{FF} = \vepsilon_{TID}^T T_{zs}^T V^{-1} (I - HS) T_{zs} \vepsilon_{TID}. $$ (40)

Let $V_{MAX,FF}^2$ be the largest eigenvalue of the symmetric matrix $T_{zs}^T V^{-1} (I - HS) T_{zs}$. A conservative value for the non-centrality parameter $\lambda^2_{FF}$ of $R_c$ is:

$$ \lambda^2_{FF,MAX} = \beta_{TID} V_{MAX,FF}^2. $$ (41)

The final step of this conservative approach is to minimize the impact of $\vepsilon_{TID}$ on the residual $r$. The non-centrality parameter $\lambda^2_{SSF}$ of $\left\| r \right\|_2$ in (17) becomes:

$$ \lambda^2_{SSF} = (f + T_{zs} \vepsilon_{TID})^T V^{-1} (I - HS) (f + T_{zs} \vepsilon_{TID}) $$ (42)

The parameter $\lambda^2_{SSF}$ is the squared norm of the vector $(f + T_{zs} \vepsilon_{TID})$ weighted by the matrix $V^{-1} (I - HS)$. A lower limit on this norm is the difference between the weighted norms of $f$ and $T_{zs} \vepsilon_{TID}$, so that:

$$ \lambda_{SSF} \geq \sqrt{f^T V^{-1} (I - HS) f} - \sqrt{\vepsilon_{TID}^T T_{zs}^T V^{-1} (I - HS) T_{zs} \vepsilon_{TID}}. $$ (43)

An upper bound on the term $\vepsilon_{TID}^T T_{zs}^T V^{-1} (I - HS) T_{zs} \vepsilon_{TID}$ was computed in (41), so that (43) becomes:

$$ \lambda_{SSF} \geq \sqrt{f^T V^{-1} (I - HS) f} - \lambda_{FF,MAX}. $$ (44)

Therefore, a conservative value for the residual’s non-centrality
parameter is:

\[ \lambda_{\text{SSF},\text{MIN}} = \sqrt{f^2 V^{-1} (1-HS)f - \lambda_{\text{RF},\text{MAX}}} \]  

(45)

Equations (39), (41) and (45) are used in Sec VII to compute the probability \( P_{sf} \) and hence, to evaluate the SSF performance in the presence of mis-modeling errors.

III. IONOSPHERIC ERROR MODELS

Ionospheric errors affecting satellite ranging measurements can be effectively eliminated using dual-frequency signals or differential corrections from nearby reference stations. In this work, an attempt is made at fulfilling stringent integrity requirements using a non-differential single-frequency navigation system configuration. To achieve this objective, reference [4] introduced a nominal measurement error model of single-frequency signals from both GPS and LEO satellites, which is refined in this section.

A. Traveling Ionospheric Disturbances (TIDs) Properties

The motivation for designing a new ionospheric error model stems from observations of experimental data. Fig. 2 is a preview of ionospheric delay measurements analyzed in Sec. IV. These measurements are derived from dual-frequency GPS carrier phase data (processing detailed in Sec. IV) collected at the Battle Creek, Michigan CORS reference station on January 8, 2007, which was a day of quiet ionospheric activity (A\(\text{K} \) index of 2 [15]). Between 2:26pm and 2:48pm local time, the PRN 13 satellite elevation angle was increasing between 10 deg and 14 deg. Over this 22min-long time span, the satellite signal travelled a distance \( d_{\text{pp}} \) of 700km across the ionosphere, at the height \( h_0 \) of 350km where the highest electron density is assumed (\( d_{\text{pp}} \) computations are explained in Sec III-B). The upper graph displays both the measured ionospheric error (thick line) and the error estimated using a piecewise linear model of the vertical ionospheric delay described in Sec. III-B (thin dashed line). This error model is used to detrend the data in the lower graph (obtained by differencing the thick and thin lines of the upper plot). The model accounts for the largest part of the ionospheric delay, but significant residual mis-modeling errors remain.

The wave-like structure observed in Fig. 2 is actually a travelling ionospheric disturbance (TID). Evidence corroborating this statement is provided in Sec V-B with an experimental analysis of the ionosphere for data collected over several months at multiple locations.

TIDs are quasi-periodic propagating ionospheric disturbances. As a subject of research over several decades, they have been observed at all latitudes, longitudes, times of day, season, and solar cycle [16] [17]. They are generally thought to be caused by the propagation of atmospheric gravity waves through the neutral atmosphere, which in turn perturbs the ionospheric electron density distribution.

![Figure 2. Observation of a Wave-Like Structure in Ionospheric Delay Measurements](image)

Typical values for the velocity, wavelength and amplitude of medium-scale TIDs are respectively 50-300m/s [18, 19, 20], 100-300km [5, 21, 22], and up to 3.5 total electron content units (TECU) [23, 24], which is equivalent to 0.57m of delay at L1 frequency. Higher values for all three quantities are found for large-scale TIDs [25, 26, 27]. Using GPS data from dense networks of ground receivers, TID wave-front widths of 2000km were measured [27], [28]. The wide range of values for TID wavelength, velocity, and amplitude highlight the unpredictability of this source of satellite ranging error.

B. Nominal Model for the Main Trend of the Ionospheric Delay

A piecewise linear model of the vertical ionospheric delay under anomaly-free conditions in mid-latitude regions was derived in [4]. The model hinges on three major assumptions.

- The ionosphere is assumed constant over short periods of times in a geocentric solar magnetospheric (GSM) frame (whose x-axis points toward the sun and whose z-axis is the projection of the earth's magnetic dipole axis on to the plane perpendicular to the x-axis) [29].

- A spherical thin shell approximation is adopted to localize the effect of the ionosphere at the altitude \( h_0 \) of the expected peak in electron density (\( h_0 = 350 \)km). An ionospheric pierce point (IPP) is defined as the intersection between the satellite LOS and the thin shell.

- The vertical ionospheric delay varies linearly with IPP separation distances \( d_{\text{pp}} \) (actually ‘great circle distances’) [30], for \( d_{\text{pp}} \) smaller than a limit \( d_{\text{pp,MAX}} \) of 750 km [4]. The distribution of the corresponding slope can be bounded by a Gaussian model [31].

Thus, the ionospheric delay for satellite \( s \) at epoch \( k \) is modeled as an initial vertical ionospheric bias \( b_{\text{ssf}} \) associated
with a ramp, whose constant slope over $d_{IPP,k}$ is the vertical ionospheric gradient $g_{TT}$. An obliquity factor $'c_{OT,k}$ accounts for the fact that the LOS pierces the ionosphere with a slant angle function of the satellite elevation angle $'\theta_k$ (e.g., [32]). The coefficient $'c_{OT,k}$ is expressed as:

$$'c_{OT,k} = \left(1 - \left[R_e \cos '\theta_k /(R_e + h_I)\right]^2\right)^{-1/2},$$  \hspace{1cm}(46)$$

where $R_e$ is the radius of the earth. As a result, the slant ionospheric delay (for code) or advance (for carrier) is given by:

$$'\varepsilon_{t,k} = 'c_{OT,k} \left( 'h_I + d_{IPP,k} 'g_{TT}\right).$$  \hspace{1cm}(47)$$

Equation (47) is applied to 750 km-long segments of IPP displacements (see [4] for details). It assumes that IPPs follow straight paths along the great circle over short time periods. The model is applicable to both GPS and LEO SV signals because $d_{IPP,k}$ is evaluated in a GSM frame fixed to the sun, the source of the ionosphere’s main variations (the motion of the earth is implicitly accounted for). Bounds on the prior probability distributions of $h_I$ and $g_{TT}$ are given by (using the notation defined in (16)):

$$h_I \sim N(0, \sigma_{h_I}^2) \text{ and } g_{TT} \sim N(0, \sigma_{g_{TT}}^2).$$  \hspace{1cm}(48)$$

This ionospheric error model is practical to utilize in a Kalman filter or in a least squares batch. It was evaluated against a set of ionospheric delay measurements collected between January and August 2007 at the CORS site of Holland, Michigan. The model proved to accurately match the data for satellite signal elevation angles higher than 50 deg [4]. However in Fig. 2, although the largest part of the ionospheric delay is efficiently removed, substantial mis-modeling errors remain that are caused by TIDs affecting low elevation measurements. A new model is derived in an attempt to mitigate the impact of these TIDs.

C. New Ionospheric Error Model Accounting for TIDs

The experimental evaluation in Sec. IV-D establishes that a substantial percentage of GPS satellite passes are affected by TIDs. Therefore, TIDs can not be considered rare-event faults for high-integrity applications and must be part of the nominal FF model.

The new model for the ionospheric delay is the sum of the nominal model in (47), which captures the main trend of the ionosphere, and a sinusoidal wave, which accounts for TIDs. The sine-wave’s amplitude is modulated by the obliquity coefficient $'c_{OT,k}$. The proposed model is written as:

$$'\varepsilon_{t,k} = C_{OT,k} \left[ h_I + d_{IPP,k} g_{TT} \right]$$

$$\left[ a_c \cos (\omega_t d_{IPP,k} + \phi) + a_s \sin (\omega_t d_{IPP,k}) \right]$$  \hspace{1cm}(49)$$

where $\omega_t$ is the TID frequency (in radians per length of IPP displacement), and $a_c$ and $a_s$ are the amplitudes of the TID’s cosine and sine terms in the vertical direction (i.e., perpendicularly to the ionospheric thin shell). The sum of sine and cosine terms in (49) is equivalent to a single cosine term with amplitude $a$ and phase $\phi$:

$$a \cos (\omega_t d_{IPP,k} + \phi),$$  \hspace{1cm}(50)$$

The expression in (49) is preferred because it only includes one non-linear parameter $\omega_t$ for the vertical delay versus two in (50) with $\omega_t$ and $\phi$.

IV. EXPERIMENTAL IONOSPHERIC MODEL ANALYSIS

Dual-frequency GPS data are used in this section to analyze ranging errors caused by the ionosphere. Processing methods implemented to fit the nominal and new error models to the data are described in Sec. IV-A, B and C. These methods are employed in Sec. IV-D to analyze TIDs. The two ionospheric error models are then evaluated in Sec. IV-E over finite windows in time (which reflects the impact on GPS signals), and in Sec. IV-F over limited IPP displacements (to measure the effect on LEO observations).

The set of GPS data selected in this analysis was collected over 91 days of quiet ionospheric activity between January and August 2007 ($A_k$-indexes ranging between 1 and 3 [15]), at seven different CORS sites spread across the United States (listed in Table I). The choice of quiet days in 2007, a year of low activity in the 11 year-long solar cycle, will provide rather optimistic results. So, if the models do not match the data for quiet days, new approaches will have to be adopted.

<table>
<thead>
<tr>
<th>Site Location</th>
<th>Latitude (deg N)</th>
<th>Longitude (deg E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battle Creek, Michigan</td>
<td>43.31</td>
<td>-86.20</td>
</tr>
<tr>
<td>Holland, Michigan</td>
<td>42.79</td>
<td>-86.11</td>
</tr>
<tr>
<td>Cleveland, Ohio</td>
<td>41.48</td>
<td>-81.67</td>
</tr>
<tr>
<td>Miami, Florida</td>
<td>25.78</td>
<td>-80.22</td>
</tr>
<tr>
<td>Houston, Texas</td>
<td>29.76</td>
<td>-95.38</td>
</tr>
<tr>
<td>Los Angeles, California</td>
<td>34.05</td>
<td>-118.25</td>
</tr>
<tr>
<td>Salt Lake City, Utah</td>
<td>40.75</td>
<td>-111.88</td>
</tr>
</tbody>
</table>
A. Experimental Data Processing Method

The ionospheric delay is proportional to the total electron content in the path of the signal and to the inverse square of the carrier frequency. This frequency-dependence is exploited here with dual-frequency satellite signals to measure ionospheric disturbances.

Dual-frequency GPS carrier-phase observations (noted $\phi_{t1}$ and $\phi_{t2}$, in units of meters, at frequencies $f_{L1} = 1575$MHz and $f_{L2} = 1228$MHz) are processed to evaluate the ionospheric delay on L1 signals (at $f_{L1}$ frequency). The following frequency coefficient is used in the derivation:

$$c_{L1} = \frac{f_{L1}^2 - f_{L2}^2}{f_{L1} - f_{L2}}$$

A biased and noisy measure of the ionospheric delay is obtained by differencing L1 and L2 observations (see for example [32]):

$$z_{t,k} = c_{L1}(\phi_{t1,k} - \phi_{t2,k}) = e_{t,k} + b_{t} + v_{t,k},$$

(51)

where $b_{t}$ is a constant bias that includes the differenced L1-L2 cycle ambiguity and inter-frequency biases. The measurement noise $v_{t,k}$ is a time-correlated random sequence such that:

$$v_{t,k} \sim N(0, \sigma^2_t),$$

where the standard deviation $\sigma_t$ is conservatively estimated to be 2.3cm (based on experimental CORS data).

Models of the ionospheric delay are evaluated using the measurement in (51). The residual errors, after removing the estimated delay from the actual data, are the sum of the measurement error $v_{t,k}$ and the modeling error.

B. Nominal Model Parameter Estimation

A batch least-squares measurement equation is obtained by substituting (47) into (51) and stacking observations $z_{t,k}$ (detrended using the Klobuchar model [33]) in order to simultaneously estimate the constant parameters $b_{t}$, $b_{t}$, and $g_{t}$:

$$
\begin{bmatrix}
  z_{1,0} \\
  \vdots \\
  z_{k,k}
\end{bmatrix} = 
\begin{bmatrix}
  1 & c_{0L,0} & c_{0L,dPP,0} \\
  \vdots & \vdots & \vdots \\
  1 & c_{0L,k} & c_{0L,dPP,k}
\end{bmatrix} 
\begin{bmatrix}
  b_{t} \\
  b_{t} \\
  g_{t}
\end{bmatrix} + 
\begin{bmatrix}
  v_{1,0} \\
  \vdots \\
  v_{k,k}
\end{bmatrix}.
$$

(52)

Equation (52) is in the form of (4). In this paper, time-correlated Gauss Markov process with time constant $T_{ig}$ is considered. The measurement noise vector $v$ in (4). The time-correlation between two measurements at sample times $t_i$ and $t_j$ is modeled in $(i,j)$-elements of $V$ as

$$2c_{L1}^2\sigma^2_{0L,\phi}e^{-\Delta t_{ij}/T_{ig}},$$

where $\Delta t_{ij} = |t_i - t_j|$. The quantity

$$2c_{L1}^2\sigma^2_{0L,\phi}$$

is also added to the diagonal elements of $V$ to account for uncorrelated carrier phase receiver noise. Prior knowledge on $b_{t}$ and $g_{t}$ is introduced assuming a-priori bounding values for their standard deviations.

C. New Model Parameter Estimation

Substituting the new model (49) into (51) establishes a non-linear relationship between the measurement $z_{t,k}$ and a constant state vector $x_{t}$ to be estimated, which is noted:

$$z_{i,k} = f_{t}(x_t) + v_{i,k},$$

(53)

where

$$x_t = [b_t, b_t, g_t, a_C, a_S, \omega_t]^T.$$

(54)

Initial guesses of the states $b_{t,0}$, $b_{t,0}$, $g_{t,0}$, $a_{C,0}$, $a_{S,0}$, and $\omega_{t,0}$ (listed in Table II) are iteratively refined using the Newton Raphson method [6]. At the $j$th iteration, (53) can be linearized about estimated parameter values $b_{t,j}$, $b_{t,j}$, $g_{t,j}$, $a_{C,j}$, $a_{S,j}$, and $\omega_{t,j}$ that are arranged in a vector $x_{t,j}$ akin to (54). The vector $\delta x_{t,j}$ of deviations between the true and estimated state parameters is defined as:

$$\delta x_{t,j} = [\delta b_t, \delta b_t, \delta g_t, \delta a_C, \delta a_S, \delta \omega_t]^T.$$
The linearized measurement equation becomes:

$$\delta z_{t,k,j} = \nabla e^T_{t,k,j} \delta x_{t,j} + v_{t,k} ,$$

where

$$\delta z_{t,k,j} = z_{t,k} - f_t(x_{t,j})$$

and

$$\nabla e_{t,k,j} = \begin{bmatrix} \frac{\partial e_{t,k,j}}{\partial b_{yt}} & \frac{\partial e_{t,k,j}}{\partial g_{yt}} & \frac{\partial e_{t,k,j}}{\partial a_c} & \frac{\partial e_{t,k,j}}{\partial a_s} & \frac{\partial e_{t,k,j}}{\partial a_t} \end{bmatrix}^T$$

with

$$\frac{\partial e_{t,k,j}}{\partial b_{yt}} = c_{ot,k}, \quad \frac{\partial e_{t,k,j}}{\partial g_{yt}} = c_{ot,k}d_{pp,k},$$

$$\frac{\partial e_{t,k,j}}{\partial a_c} = c_{ot,k} \cos(\omega_{t,j}d_{pp,k}),$$

$$\frac{\partial e_{t,k,j}}{\partial a_s} = c_{ot,k} \sin(\omega_{t,j}d_{pp,k}),$$

$$\frac{\partial e_{t,k,j}}{\partial a_t} = c_{ot,k}d_{pp,k}a_{c,j} \cos(\omega_{t,j}d_{pp,k}) - c_{ot,k}d_{pp,k}a_{s,j} \sin(\omega_{t,j}d_{pp,k})$$

Measurement deviations $\delta z_{t,k,j}$ are stacked in a batch:

$$\begin{bmatrix} \delta z_{t,1,0} \\ \vdots \\ \delta z_{t,1,kv} \\ \delta z_{t,2,0} \\ \vdots \\ \delta z_{t,2,kv} \\ \vdots \\ \delta z_{t,n,0} \\ \vdots \\ \delta z_{t,n,kv} \end{bmatrix} = \begin{bmatrix} \nabla e^T_{t,1,0} \\ \vdots \\ \nabla e^T_{t,1,kv} \\ \vdots \\ \nabla e^T_{t,n,0} \\ \vdots \\ \nabla e^T_{t,n,kv} \end{bmatrix} \delta x_{t,j} + \begin{bmatrix} v_{t,1,0} \\ \vdots \\ v_{t,1,kv} \\ \vdots \\ v_{t,n,0} \\ \vdots \\ v_{t,n,kv} \end{bmatrix} .$$

The weighted pseudo-inverse matrix $S_j$ of $[\nabla e_{t,1,0} \ldots \nabla e_{t,1,kv} \nabla e_{t,2,0} \ldots \nabla e_{t,2,kv} \ldots \nabla e_{t,n,0} \ldots \nabla e_{t,n,kv}]$ is again derived using (8) and (9). In this case, $p_{kk}$ is a 5x5 diagonal covariance matrix with diagonal elements $\sigma_{v_{IIO}}, \sigma_{v_{IIO}}, \sigma_{a_{c,0}}, \sigma_{a_{s,0}}, \sigma_{a_{o,0}}$ (also listed in Table II). The covariance matrix $V$ of the measurement vector is identical to Sec. IV-B. The state estimate vector $x_{t,j}$ is updated following the equation:

$$x_{t,j} = x_{t,j-1} + \delta x_{t,j}$$

where

$$\delta x_{t,j} = S_j \delta z_{t,j} .$$

Batch measurement updates are repeated until the convergence criterion is met:

$$|J_{j+1} - J_j| < \xi$$

where

$$J_j = \delta z_{t,j}^T V^{-1} \delta z_{t,j} .$$

A value of 0.007 was allocated to $\xi$ so that the process output accurate matches between the model and the data. If the convergence criterion in (59) is not met within a limited number of iterations, the estimation process is stopped, and the model for this sequence of measurements is categorized as non-converging.

D. Experimental TID Analysis

The first step in this experimental investigation is to confirm that the ionosphere is the cause of wave-like structures affecting sequences of observations $z_{t,k}$ in (51) (one example is identified in Fig. 2). Time-correlated measurement error due to multipath is excluded as the source of disturbance because waves do not consistently appear on daily-repeated SV trajectories. In addition, their amplitudes and wavelengths exceed 10 cm and 10 min respectively, which are too high to be attributed to multipath error [32]. Waves were simultaneously observed from nearby sites (e.g., Battle Creek and Holland, Michigan) equipped with different antennas and receivers, which excludes ground equipment malfunctions. They were also found to affect signals from different satellites crossing same sections of the sky. All these clues strongly suggest that the ionosphere is the source of wave-like structures.

Individual profiles of residual errors can be analyzed. The wave-like structure observed in Fig. 2 exhibits an approximate wavelength of 350 km (computed in a GSM frame), and an amplitude of about 20 cm, which is consistent with medium-scale TIDs [26, 27]. It is worth noticing that in most GPS-based studies, elevation cutoff angles are set higher than 30deg (30deg in [27], 40 deg in [24], 45 deg in [28], 50 deg in [5], 60 deg in [25, 26]). In contrast, in our data analysis, a 10 deg elevation mask is implemented. This may explain why large TID amplitudes (exceeding 0.7 m) are observed, even during quiet days.

Finally, a large scale analysis is performed to determine the likelihood of a GPS satellite pass to be affected by a TID. The eight-month long set of data collected for all visible SVs at the seven CORS stations listed in Table I are detrended over 750 km-long IPP displacements using the nominal model (same procedure as in Fig. 2). A segment of a SV pass containing a TID is identified when one or more sample residual error exceeds a 10 cm threshold. For this set of data, 567 out of the total 57,478 segments contained a wave-like structure, which means that 0.99% of 750 km long segments were affected by TIDs.

E. Evaluation Over a Finite Time Interval

As mentioned in Sec. I, a limit on smoothing duration $T_f$ (of 10-to-15min) was fixed to ensure validity of measurement error models (including the satellite clock and orbit ephemeris errors). Over 10 min, GPS satellite IPPs travel less than 10 cm threshold. Therefore, in this subsection, satellite passes are divided into segments of 10 min duration.

Residual errors are computed using both the nominal model and the new model. The complete eight-month-long set of data (for all SVs and locations listed in Table I) is detrended over 201,039 segments of 10 min length (which represents a total of 4,221,819 samples). The folded cumulative distribution functions (CDFs) of residual errors using both models are plotted in Fig. 3 (thick lines), together with over-bounding Gaussian functions (thin lines). Folded CDFs represent a
standard CDF for residual errors that are negative and 1-CDF when they are positive.

Fig. 3 shows the improvement achieved using the new model (dashed lines) as compared to the nominal model (solid lines). The CDF for the new model is much tighter. Maximum residual values improve from 23.2 cm to 13.2 cm. The standard deviations of the over-bounding functions are 4.7 cm and 2.65 cm for the nominal and new models, respectively. This last value of 2.65 cm is close to the expected 2.3 cm measurement noise standard deviation, which suggests that mis-modeling errors using the new model are on the order of a few millimeters.

**F. Evaluation Over Limited IPP Displacements**

Ionospheric error models derived in Sec. III apply to both GPS and LEO satellites because they are expressed as a function of IPP displacements $d_{pp}$ in a GSM frame. It was shown in [4] that complete LEO satellite passes could reach $d_{pp}$-lengths of 3300 km. LEO satellite passes were divided into $d_{pp,\text{MAX}}$-long segments (with $d_{pp,\text{MAX}} = 750$ km) so that the vertical ionospheric delay linearity assumption remained valid.

In this subsection, dual-frequency data from GPS satellites are used to evaluate ionospheric errors affecting LEO satellite signals over 750 km-long segments (actual LEO data will be processed in future work). GPS satellites take 20 min to 40 min to reach IPP displacements of 750 km. Residual errors are again computed using the two ionospheric delay models. The complete data set is detrended over 57,478 segments of 750 km length (which represents a total of 3,777,406 samples).

Folded CDFs (thick lines) and bounding Gaussian functions (thin lines) of the residual errors are given in Fig. 4 for the new (dashed) and nominal (solid) models. The folded CDFs are expectedly wider in Fig. 4 than in Fig. 3 because of the two to four-times longer fit interval. The new model provides substantial improvement. The standard deviation of the over-bounding Gaussian function decreases from 8.2 cm with the nominal model to 5.7 cm. The maximum residual for the new model was 27.2 cm versus 39.8 cm for the nominal model. The segment containing this last 40 cm residual error is plotted in Fig. 5 and shows how effective the new model (dashed line) can be. For this particular detrended delay profile (recorded for PRN 20 on 01/08/2007 starting at 10:14 am local time in Cleveland, Ohio), the maximum residual error was decreased by a factor of four.

However, there are a few cases (168 cases representing 0.29% of all segments) where the Newton-Raphson convergence criterion in (59) could not be fulfilled. The residual error profile for one of these non-converging cases (PRN 10 on 01/22/2007 at 11:11 am local time in Salt Lake City, Utah) is displayed in Fig. 6, where the data was detrended using the nominal model. The observed structure is not sinusoidal and its peak-to-valley variation exceeds 35 cm. This is evidence of the unpredictability of ionospheric disturbances, even in mid-latitude regions, during quiet days and at a low point in the solar cycle activity.

Sections III and IV have demonstrated that ionospheric delay variations are extremely challenging to capture through simple modeling. The new model brings about significant improvement, but substantial mis-modeling errors remain. The impact of these residual errors on overall system performance is evaluated in Sec. VII.
V. ORBIT EPHEMERIS ERROR MODEL

The satellite orbit ephemeris error was modeled in [3] as a constant bias (at the time the satellite first came in sight) plus a ramp over time with an unknown but constant gradient, accounting for linear variations from the initial value over the filtering interval \( T_F \). This model was adopted because GPS orbit ephemeris errors show periodic variations with frequency \( 2\pi/T_{GPS} \), where \( T_{GPS} \) is the GPS satellite orbital period of one half sidereal day. A worst slope approximation was employed because the period \( T_{GPS} \) of these variations is much larger than the interval \( T_F \) (of 10 min). The same assumption was made for LEO satellites, whose orbital period \( (T_{LAT} = 6028s) \) is much lower than \( T_{GPS} \) but still much larger than \( T_F \). The model hinged on the assumption that there is little difference between the satellite’s radial direction and the LOS to the user (only radial errors were accounted for). This approximation is reasonable for MEO spacecraft, but it is not for LEO satellites.

In response, in this section, the contributions of both radial and non-radial components of the orbit ephemeris error are evaluated for MEO and LEO spacecraft. Let us first define the nadir angle \( \alpha \). It is the angle of projection of the radial error component on the LOS direction. It is expressed as:

\[
\alpha = \sin^{-1} \left( \frac{R_e \cos \theta}{R_e + h_{SV}} \right),
\]

where, as illustrated in Fig. 7 for a simplified two-dimensional (2D) model, \( R_e \), \( h_{SV} \) and \( \theta \) respectively designate the radius of the earth, the satellite altitude and the elevation angle. The total satellite orbit ephemeris error is expressed as a function of the radial and non-radial components \( (\varepsilon_{O,R} \text{ and } \varepsilon_{O,NR} \text{, respectively})\):

\[
\varepsilon_O = \varepsilon_{O,R} \cos \alpha + \varepsilon_{O,NR} \sin \alpha.
\]

For this simplified 2D model, the coefficients of the radial and non-radial components \( (\cos \alpha \text{ and } \sin \alpha \text{, respectively})\) are plotted in Fig. 8 versus elevation angle \( \theta \). Fig. 8 shows that for MEO satellites (assuming \( h_{SV} = 20,000 \text{ km} \)), the coefficient of the radial component is much larger than the coefficient of the non-radial component for all elevation angles. In contrast with LEO satellites \( (h_{SV} = 780 \text{ km}) \), the relative values of the two components depend on satellite elevation \( \theta \).

A complete three-dimensional model is established to upgrade the satellite orbit ephemeris error model for both MEO and LEO spacecraft:

\[
{\varepsilon}^{'}_{O,A} = {\varepsilon}^{'}_{O,A} \begin{bmatrix} \varepsilon_{OM} & \varepsilon_{OC} & \varepsilon_{OR} \end{bmatrix}^T,
\]

where \( \varepsilon_{OM}, \varepsilon_{OC} \text{ and } \varepsilon_{OR} \) are the along-track, cross-track and radial orbit error components and \( {\varepsilon}^{'}_{O,A} \) is the unit LOS vector expressed in the local frame of satellite \( s \).
Orbit error variations are modeled over a time period $\Delta t_k$ from filter initiation at $t_0$ to the sample time of interest $t_k$. Over a period $\Delta t_k$ shorter than $T_p$, all three orbit error components are assumed to vary linearly:

$$\begin{bmatrix} sE_{OL,k} \\ sE_{OC,k} \\ sE_{OR,k} \end{bmatrix} = \begin{bmatrix} s b_{OL} \\ s b_{OC} \\ s b_{OR} \end{bmatrix} + \Delta t_k \begin{bmatrix} s g_{OL} \\ s g_{OC} \\ s g_{OR} \end{bmatrix} = s b_o + \Delta t_k s g_o, \quad (63)$$

The biases ($s b_{OL}$, $s b_{OC}$, $s b_{OR}$) and gradients ($s g_{OL}$, $s g_{OC}$, $s g_{OR}$) are constant, random parameters whose distributions are assumed bounded by zero mean Gaussian functions with variances $\sigma_{bo}^2$ and $\sigma_{go}^2$, respectively. Finally, back to (62), the satellite orbit ephemeris error model is expressed as:

$$s E_{OL,k} = s E_{OL,k}^T s b_o + \Delta t_k s E_{OL,k}^T s g_o. \quad (64)$$

VI. FRAMEWORK FOR THE PERFORMANCE ANALYSIS

A. Notional Space, Ground and User Segments

Three satellite constellations are considered.

- The assumed Galileo constellation is composed of 27 MEO satellites arranged in 3 regularly separated orbital planes of 9 regularly spaced spacecraft each, with a 56 deg inclination angle, and with an orbital period $T_{GAL}$ of 14 hrs, 4 min and 42s.

Angular LOS velocities are about 30 times greater for Iridium than for the MEO satellites, which in this research is a key to accomplishing rapid and robust cycle ambiguity estimation.

A notional ground segment consisting of a network of ground reference stations is considered. Ground stations are assumed co-located with WAAS reference stations, whose correction accuracy has been documented over the past seven years [37]. In the proposed architecture, Iridium satellite position and time synchronization information, together with WAAS-like ionospheric delay estimates and long-term GPS satellite error corrections, are derived at a master station using dual-frequency measurements collected at ground reference stations, and broadcast to the user via Iridium communication channels.

The conceptual navigation system described in this work is intended for civilian users. Therefore, near future implementations are based on single-frequency L-band code and carrier ranging observations. In the perspective of GPS modernization and with the emergence of new GNSS, future dual-frequency GPS and Galileo measurements are considered. Dual-frequency Iridium signals are simulated as well.

B. Complete Measurement Equation

Under nominal FF conditions (and for now, leaving the mis-modeling errors of Sec. II-E aside), the linearized carrier phase observation for a satellite $s$ at epoch $k$ is expressed as:

$$s \phi_{L,k} = s \phi_{u,k} + s N - s \phi_{T,k} + s \phi_{R,k} = s \phi_{u,k} + s \phi_{T,k} + s \phi_{M-\phi,k} + s \phi_{ES-\phi,k}, \quad (65)$$

where:

- $s u_k$ is the vector of user position (e.g., in a local reference frame) and receiver clock bias,

- $s \phi_{L,k} = \begin{bmatrix} s \phi_{u,k} \\ s \phi_{T,k} \end{bmatrix}$ and $s \phi_{L,k}$ is the unit LOS vector

- $s N$ is the carrier phase cycle ambiguity and all the other terms are measurement errors defined below.

Models were defined in previous sections for the nominal ionospheric delay $s \phi_{L,k}$ in (47) and for the satellite orbit ephemeris error $s \phi_{R,k}$ in (65). The satellite clock error $s \phi_{C,k}$ is modeled as a bias $s b_c$ plus a ramp over time $\Delta t_k$ of constant slope $s g_c$, so that:

$$s \phi_{C,k} = s b_c + \Delta t_k s g_c. \quad (66)$$
The residual tropospheric delay \( \varepsilon_{T,\Delta} \) is expressed as a zenith tropospheric bias \( b_{zT} \) plus a term that is a function of the local air refractivity index \( \Delta n \), of the difference in height \( \Delta h \), that the user experiences over the smoothing interval \( T_p \), and of the tropospheric scale height \( h_0 \), which is assigned a fixed value of 15km [38]. The tropospheric error \( \varepsilon_{T,\Delta} \) is defined as:

\[
\varepsilon_{T,\Delta} = \varepsilon_{OT,\Delta} \left( b_{zT} + 10^{-6} h_0 (1 - e^{-\Delta h / h_0}) \cdot \Delta n \right),
\]

where an obliquity factor \( \varepsilon_{OT,\Delta} \) \((\varepsilon_{OT,\Delta} = (0.002 + \sin^2(\theta))^{1/2} [38]) \) is included because a larger segment of the signal’s path travels through the troposphere at lower elevations. The parameters \( b_{zT} \) and \( \Delta n \) are not satellite-dependent because they characterize the environment surrounding the user antenna.

Parameters that are elements of the error state vector \( \mathbf{x}_{ERR} \) in (5) are summarized in the left hand side of Table III. These parameters are assumed to be constant over the smoothing period \( T_p \), and bounded by zero-mean Gaussian functions whose standard deviations (after corrections from a WAAS-like network of reference stations) are also listed in Table III and justified in [4]. These probability bounds are included in the estimation algorithm through the prior information matrix \( P_{FK}^{-1} \) (details below).

The carrier phase receiver noise \( \varepsilon_{RN,\phi,\Delta} \) in (65) is modeled as a zero-mean Gaussian white noise sequence with variance \( \sigma_{RN,\phi}^2 \). The multipath error is modeled as a first-order Gauss-Markov process with time constant \( T_M \), variance \( \sigma_{M,\phi}^2 \), and driving noise \( \nu_{M,\phi} \):

\[
\varepsilon_{M,\phi,\Delta + 1} = e^{T_M / T_u} \varepsilon_{M,\phi,\Delta} + \nu_{M,\phi}, \quad (68)
\]

with

\[

\varepsilon_{RN,\phi,\Delta} \sim N(0, \sigma_{RN,\phi}^2 (1 - e^{-2T_M / T_u})), \quad (69)
\]

The sampling interval \( T_{SB} \) within the batch of stored measurements may be different from the positioning update period.

The assertion that error models are conservative requires that the Gaussian models over-bound the CDFs of each error sources’ nominal ranging errors [39]. Alternatively, parameter values in Table III may be considered as requirements that ground corrections and user equipment should meet in order to achieve the desired system performance.

The equation for the linearized code phase measurement \( \delta \rho_{L,\Delta} \) is identical except for the absence of cycle ambiguity bias \( \nu_{N,\Delta} \) , a positive sign on the ionospheric error, and the code receiver noise \( \varepsilon_{RN,\rho,\Delta} \) (with \( \varepsilon_{RN,\rho,\Delta} \sim N(0, \sigma_{RN,\rho}^2) \)) and multipath \( \varepsilon_{M,\rho,\Delta} \) (with, in \( (69): \varepsilon_{RN,\rho,\Delta} \sim N(0, \sigma_{RN,\rho}^2 (1 - e^{-2T_M / T_u}))) \) which replace \( \varepsilon_{RN,\phi,\Delta} \) and \( \varepsilon_{M,\phi,\Delta} \), respectively.

Code and carrier phase GPS and Iridium measurements are stacked together following (1) to (3). The error state vector \( \mathbf{x}_{ERR} \) is expressed as:

\[
\mathbf{x}_{ERR} = \begin{bmatrix} \mathbf{b}_{vT}^T & \mathbf{g}_{OT}^T & \mathbf{b}_{c}^T & \mathbf{g}_{C}^T & \mathbf{b}_{zT} & \Delta n \end{bmatrix}^T, \quad (70)
\]

where individual error state vectors \( \mathbf{b}_{vT} \), \( \mathbf{g}_{OT} \), \( \mathbf{b}_{c} \), \( \mathbf{g}_{C} \) include the error states for all \( n_S \) visible satellites as described in (6), and \( \mathbf{b}_{c} \) and \( \mathbf{g}_{C} \) in (64) are constructed following:

\[
\mathbf{b}_{c} = \begin{bmatrix} \mathbf{b}_c^T & \ldots & \mathbf{b}_c^T \end{bmatrix}^T \quad \text{and} \quad \mathbf{g}_{C} = \begin{bmatrix} \mathbf{g}_c^T & \ldots & \mathbf{g}_c^T \end{bmatrix}^T. \quad (71)
\]

The prior knowledge matrix \( P_{FK}^{-1} \) on \( \mathbf{x}_{ERR} \) is diagonal, with diagonal vector:

\[
\begin{bmatrix} \mathbf{1}_{1 \times \sigma_{vT}^2} \mathbf{1}_{1 \times \sigma_{OT}^2} \mathbf{1}_{1 \times \sigma_{OT}^2} \mathbf{1}_{1 \times \sigma_{CB}^2} \mathbf{1}_{1 \times \sigma_{CG}^2} \ldots \mathbf{1}_{1 \times \sigma_{ZTD}^2} \mathbf{1}_{1 \times \sigma_{M}^2} \end{bmatrix}.
\]

Table III contains the nominal values of error parameters.

<table>
<thead>
<tr>
<th>Error State</th>
<th>Standard Deviation</th>
<th>A-priori Value</th>
<th>Other Parameters</th>
<th>Nominal Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_{vT} )</td>
<td>( \sigma_{vT} )</td>
<td>1.5 m</td>
<td>( d_{PR, MAX} )</td>
<td>750km</td>
</tr>
<tr>
<td>( \sigma_{vT} )</td>
<td></td>
<td>4 mm/km</td>
<td>( \sigma_{RN,\rho} )</td>
<td>0.3m</td>
</tr>
<tr>
<td>( \sigma_{vT} )</td>
<td></td>
<td>1.1 m</td>
<td>( \sigma_{M,\phi} )</td>
<td>3mm</td>
</tr>
<tr>
<td>( \sigma_{vT} )</td>
<td></td>
<td>0.1 m</td>
<td>( \sigma_{M,\phi} )</td>
<td>1m</td>
</tr>
<tr>
<td>( \sigma_{vT} )</td>
<td></td>
<td>1.7 m</td>
<td>( \sigma_{M,\phi} )</td>
<td>0.01m</td>
</tr>
<tr>
<td>( \sigma_{vT} )</td>
<td></td>
<td>4.44 \times 10^{-4} m/s</td>
<td>( T_{M, OBS} )</td>
<td>1min</td>
</tr>
<tr>
<td>( \sigma_{vT} )</td>
<td></td>
<td>4.44 \times 10^{-4} m/s</td>
<td>( T_{M, OBS} )</td>
<td>2s</td>
</tr>
<tr>
<td>( \sigma_{vT} )</td>
<td></td>
<td>0.12 m</td>
<td>( T_{M, OBS} )</td>
<td>30s</td>
</tr>
</tbody>
</table>

For dual-frequency: VIG and VB terms are eliminated. **For dual-frequency (at \( f_1 \) and \( f_2 \)), these terms are multiplied by \( (f_1^2 / (f_1^2 - f_2^2))^2 + (f_2 / (f_1^2 - f_2^2))^2 \)**.
C. Benchmark Mission and Requirement Definition

A benchmark mission of aircraft precision approach is assumed for performance evaluations in Sec. VII. During an approach, the airplane is assumed to follow a straight-in trajectory at a constant speed of 70m/s with a 3 deg glide-slope angle towards the runway until touchdown (TD) where requirements apply. The Miami location is selected, because, due to Iridium’s near-polar orbits, satellite geometries at this southern latitude are some of the poorest for the contiguous United States (CONUS). A nominal smoothing period $T_f$ of 10 min is chosen to investigate performance variations.

Three fundamental navigation performance metrics, originating from safety-critical aviation applications [38], are emphasized in the integrity analysis. The first metric is integrity, defined as the ability of a system to provide timely warnings in case of hazardous navigation error. The integrity requirement $P_{inn}$ or probability of hazardous misleading information (HMI) is given a value of $2 \cdot 10^{-7}$ [38]. Second, continuity is the likelihood that the system meets the integrity requirement over the entire mission duration, with no unscheduled interruption. A value of $8 \cdot 10^{-6}$ is considered for the continuity requirement $P_c$ [38]. Finally, availability is the fraction of time that integrity and continuity requirements are fulfilled. Integrity and availability are instantaneous measures of mission safety, whereas availability is evaluated over multiple operations.

The most harmful threats to navigation system integrity and continuity are rare-event measurement faults such as ground or user equipment malfunctions, unusual atmospheric conditions or satellite failures. User receiver malfunctions can be mitigated by avionics redundancy. Ground monitoring can be implemented to ensure the robustness of broadcast corrections. Anomalous conditions in the signal propagation environment are extremely challenging to model, and will be the subject of future investigations. Therefore, in this early stage of the system design, the integrity analysis focuses on satellite faults. Reference [40] specifies that the GPS satellite failure rate $\Lambda$ is $10^{-9}$/hr. The prior probability $P_f$ of an individual satellite fault occurring during the exposure period $T_f$ is: $P_f = \Lambda T_f$.

The overall integrity requirement $P_{inn}$ is allocated (for practical reasons in the implementation of estimation and detection algorithms) between mutually exclusive hypotheses of fault-free (FF) operation, single-satellite fault (SSF) conditions, and all other conditions as illustrated in Fig. 9. In this work, an integrity risk $P_{MF}$ is set aside for cases of multiple SV faults occurring during the same time interval $T_f$.

Multiple simultaneous faults are assumed independent events and hence have a low probability of occurrence. Therefore, the value of $P_{MF}$ can be selected larger than the probability of two or more faults occurring during $T_f$, so that:

$$P_{MF} \geq 1 - \sum_{i=0}^{n_S} C_i^{n_S} P_f^i (1 - P_f)^{n_S - i} ,$$

where $n_S$ is the number of visible SVs and $C_i^{n_S}$ is the binomial coefficient. For a 10 min exposure period $T_f$ and using measurements from ten different SVs, the probability $P_{MF}$ is on the order of $10^{-8}$. Then, an integrity budget of $\gamma(P_{inn} - P_{MF})$ is allocated to normal FF conditions, and the remaining fraction $\left(1 - \gamma\right)(P_{inn} - P_{MF})$ is attributed to SSF. The coefficient $\gamma$ ranges between 0 and 1; a value of 0.01 is selected to maximize the combined FF-SSF performance defined below.

Availability performance criteria are established for the FF and SSF hypotheses. In this work, the focus is on the vertical coordinate both because of the tighter requirements in this direction and because of the generally worse vertical positioning performance as compared to horizontal coordinates. Under normal FF conditions, the vertical protection level $VPL$ is expressed as:

$$VPL = \kappa_{FF} \sigma_u + b ,$$

where $\sigma_u$ is defined in (11), and the probability multiplier $\kappa_{FF}$ is the value for which the normal CDF equals $1 - \gamma(P_{inn} - P_{MF})/2$. In Sec. VII, the TID mis-modeling error parameter $b$ is zero unless specified otherwise, in which case it is conservatively defined by (38). In accordance with civilian aviation standards, which specify a vertical alert limit $VAL$ of 10m from 200 feet of altitude to TD [38], a satellite geometry is deemed available under FF conditions if and only if:

$$VPL < VAL .$$

In addition, SSF availability is granted if and only if:

$$P_{MD} < (1 - \gamma)(P_{inn} - P_{MF})/P_f ,$$

where the probability of missed detection $P_{MD}$ is defined in (20). A refined version of (75) is given in [4] for ramp-type faults starting at frequent intervals. Equations (74) and (75) are the expressions of FF and SSF binary criteria that either validate or nullify availability for an approach.
Of particular importance when combining measurements from multiple constellations is the duration $T_{AV}$ over which availability simulations are carried out. The period $T_{AV}$ should enable sampling of a complete set of satellite geometries. The GPS satellite geometry repeats itself every sidereal day (corresponding to $2\pi_{GPS}$). Iridium SVs circle the earth exactly 43 times in three solar days and four seconds. For the Iridium/GPS constellation, a period $T_{AV}$ of three days was shown to provide performance results representative of all geometries, while at the same time, simulating SV geometries over three days is still computationally manageable [3]. In this paper, Galileo signals are integrated. It takes ten sidereal days for Galileo satellites to circle the earth 17 times. It corresponds to an integer number of GPS orbital periods ($20\pi_{GPS}$) and to a large number of Iridium orbital periods ($142.94\pi_{IRIT}$). The 0.06$T_{IRIT}$ difference (6 min) to the next integer number of $T_{IRIT}$ averages out over ten days. Therefore, aircraft approaches starting at regular 30s intervals are simulated over a period $T_{AV}$ of ten sidereal days. The percentage of approaches that satisfy both (74) and (75) is the measure of ‘combined FF-SSF availability’ used in the integrity analysis of Sec. VII.

### VII. Multi-Constellation Availability Performance Analysis

#### A. Analysis of New RAIM-Based Methods

The limits of the detection algorithm are evaluated using worst-case fault modes that minimize the probability of missed detection $P_{MD}$. Simulations are performed for the nominal single-frequency, ground-corrected Iridium/GPS system configuration synthesized in (65). Combined FF-SSF availability results are summarized in Table IV.

Table IV shows that the performance drops from 100% when evaluated against step, ramp and impulse-type faults (of all magnitudes and start times) to 56.2% against theoretical worst-case faults. Much of the lost availability can be recovered using the extended window RAIM algorithm. Performance at the Miami location improves from 56.2% to 95.8% assuming a 15 min extended-window period $T_{EWT}$.

It is also shown that the extended window algorithm works as an efficient exclusion mechanism. The 95.8% availability number is compared to the result of a 15 min-long batch that includes all Iridium observations (even between times $t_i$ and $t_f$ in Fig. 1, i.e., $T_f = 15$ min): in this case availability only reaches 86.8%. Therefore, simply extending the Iridium/GPS batch-interval would not be an effective solution.

Finally, Miami was chosen as a near-worst case location for Iridium/GPS positioning over CONUS. In comparison, combined availability of 99.95% is achieved at Chicago. Also, the extended window RAIM method produces better results if $T_{EWT}$ is increased to 20 min: in this case, availability at Miami increases from 95.8% to 99.8%, provided that the validity of GPS measurement error models can be validated over 20 min.

#### B. Measurement Error Models: Impact of TIDs

The added uncertainty introduced by non-radial components of the satellite orbit ephemeris error in the updated model of Sec. V causes no variation in availability. The two reasons for this result are the high orbit altitude of GPS satellites (which yield a small non-radial coefficient illustrated in Fig. 8) and the small magnitude of the assumed orbit ephemeris error for Iridium SVs (small $\sigma_{OB}$ and $\sigma_{OG}$ values in Table III).

In contrast, accounting for TIDs plays a crucial part in the availability performance evaluation. As mentioned previously, TIDs are frequently-occurring, temporary, localized and moving wave-like structures that are highly unpredictable. Experimental evaluation of two ionospheric error models in Sec. IV shows that TIDs cause decimeter-level residual errors. In order to reduce mis-modeling errors, more robust ionospheric error models could be implemented. However, this approach is limited because each new parameter added to the model introduces uncertainty in the estimation algorithm, which ultimately causes loss of availability.

As an alternative, ground correction and monitoring of the ionosphere against TIDs is considered. Fig. 10 displays IPP locations (computed in a GSM frame) over a 10 min period for a network of 28 ground receivers co-located with WAAS reference stations. GPS IPP separation distances are too large to robustly detect structures at the 100-300 km scale (typical TID wavelength). The snapshot of IPP coverage in Fig. 10 demonstrates that even if Iridium signals were used, large sections of the sky remain uncovered.

Instead, in this paper, residual errors caused by TIDs are conservatively accounted for using the method derived in Sec. II-E. Vertical protection levels (VPLs) at touch down are plotted in Fig. 11 for aircraft approaches simulated over a 24 hour period, for the Iridium/GPS constellation at the Miami location (with $T_f = 10$ min). The VPL saw-tooth pattern is explained as follows [3]: low VPLs are achieved at the seam of the Iridium constellation where the orbital plane separation angle is smaller (corresponding to about 2-3 hr and 14-15 hr on the x-axis); the Miami location crosses an Iridium orbital plane every two hours, which generates regularly-spaced valleys in VPL. These observations demonstrate that Iridium satellite signals are the driving force in Iridium/GPS performance.

<table>
<thead>
<tr>
<th>Step/ramp/impulse</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worst-case faults</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Extended-window</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Miami location</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Chicago location</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>$T_f$ (min)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>$T_{EWT}$ (min)</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability (%)</td>
<td>100</td>
<td>56.2</td>
<td>95.8</td>
<td>86.8</td>
<td>99.9</td>
</tr>
</tbody>
</table>

#### TABLE IV. PERFORMANCE COMPARISON OF RAIM-BASED METHODS
FF-availability of 100% is achieved assuming no occurrence of TIDs (thick line: VPLs are lower than VAL). In parallel, VPLs are conservatively estimated assuming that single-satellite TIDs could alter ranging signals during each approach. A $\beta_{\text{TD}}$ parameter value of 10 cm is implemented in (38) to compute an upper-bound $b$ on the mis-modeling error, which causes a dramatic drop in availability from 100% to 51%. Experimental data showed that, even during days of quiet ionospheric activity, higher $\beta_{\text{TD}}$-values could be expected (residual errors up to 40 cm). Therefore, this analysis shows that TIDs heavily impact the FF performance (the combined FF-SSF availability is even lower).

In future work, less conservative methods to account for mis-modeling errors will be investigated. Also, improved ground corrections could lower the $\sigma_{\text{VBR}}$ parameter, hence alleviating the need for precise TID modeling. Finally, in the next subsections, the performance of future dual-frequency implementations is assessed.

C. Future Dual-Frequency Multi-Constellation Navigation Systems

Dual-frequency measurements are free of ionospheric errors (at the cost of a slight increase in measurement noise as indicated in Table III). The theoretical worst-case performance (computed using worst-case faults) of future dual-frequency multi-constellation navigation systems is evaluated in Fig. 12. Combined FF and SSF availability is presented for a 5 deg x 5 deg latitude-longitude grid of locations over CONUS.

Results for the GPS/Galileo system were very poor and are not represented. This highlights again that large Iridium satellite motion is instrumental in meeting a 10 m VAL requirement. Combined availability for the two Iridium-augmented systems illustrated in Fig. 12 improves at higher latitudes, where the SV density of the Iridium near-polar constellation increases. Availability ranges between 95.6% and 100% for the Iridium/GPS system, and between 99.2% and 100% for Iridium/GPS/Galileo. Enhanced performance is achieved using the extended-window RAIM procedure with $T_{\text{EW}} = 15$ min: 99.8% availability is reached at the worst-case CONUS location for the Iridium/GPS system, and 100% availability is obtained at all locations using the Iridium/GPS/Galileo constellation. Finally, both Iridium-augmented GNSS produce maximum availability if the extended-window period $T_{\text{EW}}$ is increased to 20 min.
These results were computed under the assumption that the satellite clock and orbit and tropospheric error models were robust. The next phase of the analysis is to evaluate these error models against experimental data.

VIII. CONCLUSION

The combination of redundant ranging signals from multiple satellite constellations opens the possibility for high-integrity positioning over wide areas. Multi-constellation navigation systems provide geometric diversity, which can be exploited by filtering carrier phase measurements over time.

In this paper, rather than establishing a comprehensive description of all potential system failures, fault modes that maximize the probability of missed-detection were derived to directly assess the theoretical worst-case performance. Then, based on a careful analysis of undetected faults, a new extended-window RAIM method was implemented to exclude satellite geometries for which worst-case faults are unobservable using the RAIM residual. The resulting availability performance is much improved as compared to the nominal batch residual-based RAIM method, but assumes measurement error model validity over longer time-intervals.

Ionospheric error models for both GPS and LEO satellite signals were experimentally evaluated using dual-frequency GPS data collected at multiple CORS sites over an eight-month long period. A new model was derived to account for frequently-occurring, localized, temporary wave-like structures that were demonstrated to be TIDs. These TIDs cause decimeter-level residual errors, even during days of quiet ionospheric activity. A conservative approach to account for the unknown time-behavior of residual errors showed that the overall system performance was heavily impacted by TIDs.

Overall system performance sensitivity was quantified over CONUS, for three multi-constellation navigation systems (GPS/Galileo, Iridium/GPS and Iridium/GPS/Galileo). It shows that the two dual-frequency Iridium-augmented GNSS can potentially achieve maximum availability at all CONUS locations.

APPENDIX. OBSERVATION AND MEASUREMENT NOISE
Covariance Matrices

A state space representation of code-phase observation vector \( \mathbf{z} \) is realized in (4). Let \( \mathbf{0}_{n\times m} \) be a \( n \times m \) matrix of zeros. State coefficients are arranged in matrices that are needed in later steps, so that for satellite \( s \) visible between epochs 0 and \( n_s-1 \):

\[
\begin{bmatrix}
\mathbf{g}_0^T \\
\vdots \\
\mathbf{g}_{n_s-1}^T
\end{bmatrix}, \quad
\begin{bmatrix}
\cdots \\
\mathbf{c}_0 \cdots \\
\mathbf{c}_{n_s-1} \cdots
\end{bmatrix},
\]

\[
\begin{bmatrix}
\mathbf{e}_0^T \\
\vdots \\
\mathbf{e}_{n_s-1}^T
\end{bmatrix}, \quad
\begin{bmatrix}
\mathbf{d}_{r,p}^T \\
\vdots \\
\mathbf{d}_{r,p,n_s-1}^T
\end{bmatrix},
\]

Carrier phase observations for all \( n_s \) Iridium and GPS satellites are then stacked together in (2). The carrier phase measurement equation is written in the form of:

\[
\mathbf{\phi} = \mathbf{H}_s \mathbf{x} + \mathbf{v}_s,
\]

(76)

The carrier phase observation matrix \( \mathbf{H}_s \) is constructed by blocks:

\[
\mathbf{H}_s = \begin{bmatrix}
\mathbf{G} & \mathbf{B}_N & \mathbf{B}_{VIB} & \mathbf{B}_{VIG} & \mathbf{B}_{OB} & \mathbf{B}_{OG} & \cdots & \mathbf{B}_{CB} & \mathbf{B}_{CG} & \mathbf{B}_{ZTD} & \mathbf{B}_{An}
\end{bmatrix}
\]

Each block corresponds to a vector of state parameters, and contains coefficients for all spacecraft, for the entire sequence of measurements. Let \( n_p(s) \) be the number of samples for satellite \( s \) (which generally differs for fast-moving LEO SVs), and \( \mathbf{I}_{n \times 1} \) be a \( n \times 1 \) column-vector of ones:

\[
\mathbf{G} = \begin{bmatrix}
\mathbf{1}_G \\
\vdots \\
\mathbf{1}_G
\mathbf{n}_s \mathbf{G}
\end{bmatrix}, \quad
\mathbf{B}_N = \mathbf{B}_{UB} = \begin{bmatrix}
\mathbf{1}_{n_p(1)} \mathbf{c}_0 & 0 \\
\vdots & \ddots & \vdots \\
0 & \mathbf{1}_{n_p(n_s)} \mathbf{c}_{n_s-1}
\end{bmatrix},
\]

\[
\mathbf{B}_{VIB} = \begin{bmatrix}
\mathbf{1}_{\mathbf{c}_{OL}} \mathbf{c}_{OL} & 0 \\
\vdots & \ddots & \vdots \\
0 & \mathbf{1}_{n_c} \mathbf{c}_I
\end{bmatrix}, \quad
\mathbf{B}_{VIG} = \begin{bmatrix}
\mathbf{1}_{\mathbf{c}_{OL}} \mathbf{c}_{OL} & 0 \\
\vdots & \ddots & \vdots \\
0 & \mathbf{1}_{n_c} \mathbf{c}_I
\end{bmatrix},
\]

\[
\mathbf{B}_{OB} = \begin{bmatrix}
\mathbf{1}_{\mathbf{e}_O} \mathbf{e}_O & 0 \\
\vdots & \ddots & \vdots \\
0 & \mathbf{1}_{n_c} \mathbf{e}_{n_s}
\end{bmatrix}, \quad
\mathbf{B}_{OG} = \begin{bmatrix}
\mathbf{1}_{\mathbf{e}_O} \mathbf{e}_O & 0 \\
\vdots & \ddots & \vdots \\
0 & \mathbf{1}_{n_c} \mathbf{e}_{n_s}
\end{bmatrix},
\]

\[
\mathbf{B}_{CG} = \begin{bmatrix}
\mathbf{1}_{\Delta \mathbf{t}} \Delta \mathbf{t} \mathbf{0} \\
\vdots & \ddots & \ddots \\
0 & \mathbf{1}_{n_c} \Delta \mathbf{t}
\end{bmatrix}, \quad
\mathbf{B}_{ZTD} = \begin{bmatrix}
\mathbf{1}_{\mathbf{c}_{OT}} \mathbf{c}_{OT} \\
\vdots \\
\mathbf{1}_{n_c} \mathbf{c}_{OT}
\end{bmatrix}, \quad
\mathbf{B}_{An} = \begin{bmatrix}
\mathbf{1}_{\mathbf{c}_T} \mathbf{c}_T \\
\vdots \\
\mathbf{1}_{n_c} \mathbf{c}_T
\end{bmatrix}.
\]

A measurement equation similar to (76) is established for the code-phase observation vector \( \mathbf{p} \). In this case, the sign on the ionospheric coefficients \( \mathbf{B}_{VIB} \) and \( \mathbf{B}_{VIG} \) is positive. Also, the
columns of ones in $B_N$ corresponding to the cycle ambiguity vector $N$ are replaced by zeros; this explains why state vectors $N$ and $b_c$ have to be distinguished, even though columns of $B_N$ and $B_{CB}$ are linearly dependent for carrier phase measurements.

The complete sequence of code and carrier phase signals for all satellites over the smoothing interval are included into a batch measurement vector (3), which leads to the state space measurement equation (4). The measurement noise vector $v$ is utilized to introduce the time-correlated noise due to multipath modeled as a Gauss Markov process. Its covariance $V$ is block diagonal, and each block corresponds to observations from a same SV over time. Within each block, the time-correlation between two measurements originating from a same satellite $s$ at sample times $t_i$ and $t_j$ is modeled as

$$
\sigma_{\Delta t - \sigma}^2 \cdot e^{-\Delta t_i / \sigma} ,
$$

where $\Delta t_i = t_i - t_j$. $\sigma_{\Delta t - \sigma}^2$ and $\sigma_{\Delta t - \sigma}^2$ are also added to the diagonal elements of $V$ to account respectively for code-phase and for carrier phase uncorrelated receiver noise.

REFERENCES


