ORIGINAL ARTICLE

WILEY ✧ ION®

Navigation and ionosphere characterization using high-frequency signals: Models and solution concepts

Yoav Baumgarten[1] ⓘ | Mark L. Psiaki[2] ⓘ | David L. Hysell[3]

[1] Research Associate, RD&E Division, Rafael – Advanced Defense Systems LTD., Haifa 3102102, Israel

[2] Professor and Kevin T. Crofton Faculty Chair, Kevin T. Crofton Dept. of Aerospace & Ocean Engineering, Virginia Tech, Blacksburg, VA 24061, USA (Email: mlpsiaki@vt.edu)

[3] Professor and Chair, Earth and Atmospheric Sciences, Cornell University, Ithaca, NY 14853, USA (Email: dlh37@cornell.edu)

Correspondence
Yoav Baumgarten, Research Associate, RD&E Division, Rafael – Advanced Defense Systems LTD., Haifa, 3102102, Israel.
Email: yoavba@rafael.co.il

Abstract

A navigation concept is being developed that relies on passive one-way ranging using pseudorange and beat carrier-phase measurements of high-frequency (HF) beacon signals that travel along non-line-of-sight paths via ionosphere refraction. This concept is being developed as an alternative to GNSS positioning services. If the set of signals that reaches a user receiver has sufficient geometric diversity, then the position of that receiver can be determined uniquely. Ionospheric modeling uncertainty causes errors in the deduced user position, but these errors are compensated by estimating corrections to a parametric model of the ionosphere as part of the navigation solution. A batch filter estimates the user position and corrections to an *a priori* ionosphere model. In simulated case studies involving significant errors in the *a priori* ionospheric parameters, the total positioning error is on the order of tens of meters in the horizontal plane and on the order of meters vertically.

1 | INTRODUCTION

Satellite-based navigation vulnerability to electronic jamming and spoofing is of immense concern for developers of both civil and defense systems. A broad search for alternative methods has been conducted and widely discussed in the navigation literature in the past decade. High-frequency (HF) signals propagating in the atmosphere have been used for communications and over-the-horizon radar. Signals with frequencies in the range 2–10 MHz can bounce successively off the ionosphere and the Earth to arrive at a receiver along a non-line-of-sight (NLOS) path. Such signals have been proposed for geolocation purposes, as in Huang and Reinisch (2006). The present study represents a further effort to examine the potential use of such signals as an alternative radio navigation method.

Given perfect knowledge of the electron density distribution in the ionosphere and of the number of bounces

between a transmitter and a receiver, it is possible to develop a model of the measured pseudorange, also known in the literature as range-equivalent group delay, which is the difference between a signal's reception and transmission times multiplied by the speed of light. The pseudorange depends on the unknown user receiver location and the receiver's unknown clock offset. Given four or more such pseudoranges from four or more independent transmitters with an appropriately diverse geometry, it should be possible to solve for the unknown user position and clock offset, similar to GPS.

The problem with such an approach is that the ionosphere's HF signal propagation/refraction/reflection properties are highly uncertain due to the variability of its three-dimensional electron density distribution. The approach of Huang and Reinisch (2006) and Zaalov et al. (2017) is to use ionosonde data (Juras, 1985) in order to refine a local electron density model. This local model is

then used to estimate the unknown location of a transmitter. Fridman et al. (2006) suggests an alternative procedure for local electron density modeling using GPS data inversion. Nickisch et al. (2016) further expands the work of Fridman et al. (2006) to assimilate HF near-vertical propagation time delay, HF Doppler shift, and HF angle-of-arrival measurements in ionosphere 3D electron density characterization. Performance is then evaluated based on a comparison between computed and measured angle-of-arrival values, where the reference measured values are provided by highly accurate equipment.

The present study, however, seeks to estimate simultaneously the location of an unknown receiver, its clock offset, and corrections to relevant portions of the electron density distribution. The approach taken for the estimation problem, model, and solution method involves several components. They are 1) a nominal ionosphere model, 2) estimated corrections to that model, 3) raytracing calculations for the paths of the HF signals and their observables, and 4) model inversion calculations to estimate the user receiver position, the user clock offset, and corrections to the ionosphere model. These model inversion calculations are carried out using a modified nonlinear batch least-squares solution technique.

Given a limited number of transmitters and a limited number of observables, a key question for such an approach concerns ionosphere parameter observability and the extent to which position accuracy can be attained. It has been demonstrated in Baumgarten & Psiaki (2017) that it is possible to combine *a priori* information for a parameterized model of ionosphere electron density with measured pseudoranges in order to arrive at a reasonable result. The work described in Baumgarten & Psiaki (2017) demonstrated feasibility for the joint positioning/ionosphere-characterization approach based on a simplified thin-sheet ionosphere model for propagating signals. The present paper describes a follow-up study that utilizes a 3D ionosphere electron density variation model, raytracing calculations that rely on advanced signal propagation models, and an enhanced batch-filtering algorithm. Its fundamental input data are the measured pseudoranges between an array of transmitters at known locations and the user receiver. A second type of measurement, introduced to this study, is the beat carrier phase. It counts carrier cycles over an arbitrary time interval and differences the resulting count with the expected nominal count for the transmitted signal waveform (Misra & Enge, 2011). The ionosphere model utilized in both phases of this study is a Chapman vertical profile whose three parameters vary with latitude and longitude.

The current study makes three contributions to the area of radio navigation based on bouncing HF signals. First, it develops a model of the pseudorange and beat carrier-phase measurements of multi-hop HF signal paths from known beacon transmitter locations to an unknown user receiver location. This model includes techniques for solving its nonlinear bounce conditions and for computing first-partial derivative sensitivities of the bounces and of the pseudorange and beat carrier-phase measurements with respect to the unknown user location and the unknown ionosphere parameters. Second, this study develops a batch nonlinear least-squares estimation algorithm for determining the unknown user receiver position, user receiver clock offset, and ionospheric parameter corrections. This algorithm incorporates *a priori* information about the ionosphere parameters in order to compensate for the lack of strict simultaneous observability of the location, clock offset, and ionosphere parameters. Third, the potential performance of the proposed HF navigation scheme is evaluated in a preliminary manner using data from a truth-model simulation.

The remainder of this paper is divided into six sections. Section II covers the physical and mathematical models of HF signals propagating in the ionosphere. It describes the fundamentals of the raytracing computations that this study uses to model HF signal paths. Section III presents models for the Earth and the ionosphere, including the parameterized Chapman vertical profile model. Section IV covers definitions and derivation of bounce points and their equations, ray-hops, and multi-hop ray paths. It also discusses the two measurement models. Section V develops the batch filter that estimates the quantities of interest. It starts by formulating the batch-filtering problem, and it outlines a modified Gauss-Newton method that solves the problem. Section VI presents a preliminary evaluation of the batch filter's performance and this method's potential accuracy. Section VII summarizes this study's developments.

2 | SIGNAL PROPAGATION IN THE IONOSPHERE AND RAYTRACING COMPUTATIONS

2.1 | High frequency signals and signal propagation

This study considers transmitted RF signals with carriers in the range 2 MHz–10 MHz. Transmitted signals are assumed to maintain a constant carrier frequency or to utilize a smoothed stepping pattern for altering their carrier frequencies. The latter type of signal can be useful for navigation and ionosphere correction if its beat carrier phase is well defined at the transmitter and accurately measured at the receiver. Beat carrier-phase measurements are made after a given frequency step is complete and the signal is

oscillating with a new, constant frequency. These measurements are particularly useful because the signal traverses a perturbed ray path due to the frequency change and therefore probes more of the ionosphere.

The presumed ability to measure pseudorange depends on the signal having been modulated by a binary phase-shift keying (BPSK) pseudo-random code or some similar spread-spectrum modulation. The resulting accuracy for this sort of ranging in terms of measurement noise 1 sigma is about 1 kilometer assuming a signal bandwidth of 100 KHz. Carrier-phase measurements are assumed to be derived using a stable internal oscillator and a phase-lock loop so that the expected accuracy for a range-equivalent beat carrier-phase measurement is 1 meter based on extrapolation of the fraction-of-a-cycle phase accuracy that can be resolved for L-band signals using standard signal-processing techniques.

The wave propagation mechanism that is considered in this study relies on ionospheric refraction that bends skyward-propagating radio waves back towards the Earth in a way that resembles reflection. This effect can occur for signals in the frequency range of up to 40 MHz (Hysell, 2018). HF signals traveling in the ionosphere are characterized not only by a curved trajectory shape, but also by the frequency- and path-dependent propagation speeds of their BPSK modulated code and carrier phase. Propagation speed dependence on wave frequency is known as *dispersive* wave propagation and is typical of propagation in a plasma (Cummer, 2000; Ishimaru, 2017; Juras, 1985). Further details about the behavior of electromagnetic waves as they traverse through plasma are provided in (Baumgarten, 2018).

Electromagnetic wave propagation in the birefringent, inhomogeneous, and lossy ionosphere obeys the Appleton-Hartree formula (Yeh & Liu, 1972):

$$n^2 = 1 - \frac{X}{1 - \frac{Y_T{}^2}{2(1-X)} \pm \sqrt{\frac{Y_T{}^4}{4(1-X)^2} + Y_L{}^2}}, \quad (1)$$

where

$$X \equiv \frac{\omega_p{}^2}{\omega^2}; \quad Y \equiv \frac{\Omega_e}{\omega}; \quad Y_T = Y\sin(\theta); \quad Y_L = Y\cos(\theta), \quad (2)$$

and where n is the index of refraction, θ is the angle between the magnetic field $\underline{B}$ and the wave vector $\underline{k}$, ω is the wave frequency, ω_p is the plasma frequency, and Ω_e is the electron gyrofrequency. The latter is dependent on Earth's magnetic field. The two solutions for the refractive index correspond to the ordinary (O) mode and the extraor-

dinary (X) mode that have different polarization and, consequently, different ray paths and measured observables (Baumgarten, 2018; Seybold, 2005). This formula makes it straightforward to obtain the conditions for which waves propagate, i.e., the conditions on Y, X, and θ for which $n^2 > 0$.

2.2 | Raytracing

Raytracing calculations lie at the core of this study. The ability to accurately model signal trajectories in the ionosphere is essential to the modeling of the pseudorange and beat carrier-phase observables. Calculations are based on a numerical solution to the wave equations through propagation of Hamilton's equations that apply for an RF signal traversing in an ionized medium (Bennett et al., 2004). The fundamental set of raytracing equations is provided by Jones & Stephenson (1975) in the form of nonlinear ordinary differential equations that can be written as

$$\frac{dr_{\underline{w}}}{dP'} = -\frac{\left[\frac{dH}{d\underline{k}}\right]^T}{\left[c\frac{dH}{d\omega}\right]} \equiv f_1\left(\underline{r}_w, \underline{k}\right)$$

$$\frac{d\underline{k}}{dP'} = \frac{\left[\frac{dH}{dr_{\underline{w}}}\right]^T}{\left[c\frac{dH}{d\omega}\right]} \equiv f_2\left(\underline{r}_w, \underline{k}\right). \quad (3)$$

H is the Hamiltonian, and the independent variable $P' \equiv ct_g$ is the range-equivalent group delay parameter that takes the value $P'_0 = 0$ at the beginning of the trajectory and P'_f at its end. The same Hamiltonian can be used to develop a differential equation for the range-equivalent carrier phase $P = \phi/k_0$ with ϕ being the carrier phase in radians. This differential equation takes the form:

$$\frac{dP}{dP'} = -\frac{\left[\frac{dH}{d\underline{k}}\underline{k}\right]}{\left[\frac{dH}{dk_0}k_0\right]}, \quad (4)$$

where $k_0 = \omega/c$ is the free-space wave number with ω being the transmission frequency. Note that the last two equations would have taken a somewhat different form had the position been given in geographic coordinates as in Nickisch (1988).

Jones and Stephenson (1975), which combines the work of Lighthill (1965) and Haselgrove (1954), gives several Hamiltonians that can be used in Equation (3). They are

generally based on the Appleton-Hartree formula. Alternative Hamiltonian formulations are presented in Psiaki (2019). The latter are the Hamiltonians that are utilized with this study. The first Hamiltonian, which is used where the electron density is relatively small, is given by

$$H = \frac{1}{2} \, real \left\{ \left[\left(\frac{\underline{k}^T \underline{k}}{k_0^2} \right) - n_{AH}^2 \left(\underline{r}_w, \frac{\underline{k}}{\|\underline{k}\|}, k_0, \underline{p} \right) \right] \right.$$
$$\left. \times \left[(1 - jZ)^2 - Y^2 \right] (1 - jZ) \right\}, \quad (5)$$

where $\underline{p}$ is a vector of parameters that characterizes the ionosphere electron density profile, and n_{AH} is the lossy Appleton-Hartree index of the refraction of Yeh & Liu (1972). A second Hamiltonian is used near a reflection point/spitze. This second Hamiltonian, given in Baumgarten (2018), is used because it does not experience any singularities in this vicinity.

A state space system of equations is defined for the unknown wave-front position and wave vector. It is based on the two expression of Equation (3) and takes the form:

$$\frac{d\underline{\tilde{X}}}{dP'} = \underline{\tilde{f}} \left(\underline{\tilde{X}} \right); \quad \underline{\tilde{f}} = \begin{bmatrix} \underline{f}_1 \left(\underline{\tilde{X}} \right) \\ \underline{f}_2 \left(\underline{\tilde{X}} \right) \end{bmatrix}; \quad \underline{\tilde{X}} = \begin{bmatrix} \underline{r}_w \\ \underline{k} \end{bmatrix}. \quad (6)$$

Thus, the state vector consists of the three Cartesian coordinates of the propagating wave front's position, $\underline{r}_w$, and the three components of its wave vector, $\underline{k}$. For practical reasons, the state vector that is used with the current numerical implementation is defined in the normalized form:

$$\underline{X} = \begin{bmatrix} \frac{\underline{r}_w}{P'_f} \\ \frac{\underline{k}}{k_0} \end{bmatrix}. \quad (7)$$

Normalization of the first term by P'_f and of the second term by k_0 results in a non-dimensional state vector, whose derivative with respect to the non-dimensionalized independent variable $\tau \equiv P'/P'_f$ is given by

$$\frac{d\underline{X}}{d\tau} = P'_f \frac{d\underline{X}}{dP'} \equiv \underline{f} \left(\underline{X}, \underline{p}, P'_f \right). \quad (8)$$

This nonlinear differential equation can be numerically propagated from the initial point at $\tau = 0$ to the final point at $\tau = 1$. The terminal value P'_f is unknown and must be determined as part of a two-point boundary value problem solution. Further details about the raytracing two-point boundary value problem numerical solution technique are given in Section III of Psiaki (2019).

3 | EARTH AND IONOSPHERE MODELS

Models of the Earth and the ionosphere are used to define the physical environment for the propagating signals. These models have been chosen because they satisfy the need for a reasonably realistic representation of physical phenomena and the need to limit the complexity of the models and the resulting computational effort.

3.1 | Earth geometry and magnetic field models

The Earth is modeled as a closed, continuous and smooth surface that is known as the WGS-84 ellipsoid (National Geospatial Intelligence Agency, n.d.). The implicit equation for the ellipsoid in Cartesian Earth Centered Earth Fixed (ECEF) coordinates is

$$\frac{r_1^2}{R_c^2} + \frac{r_2^2}{R_c^2} + \frac{r_3^2}{R_P^2} = 1$$
$$R_c = 6378137 \; [m]$$
$$R_P = 6356752.31425 \, [m], \quad (9)$$

where r_1, r_2, and r_3 are the Earth-fixed Cartesian coordinates of the surface point in meters.

This approach for modeling the Earth has been chosen for its relative simplicity and the fact that it does not rely on availability of additional data. A more realistic method for modeling the shape of the Earth would use biquintic surfaces, approximating an existing digital representation of the Earth such as a digital terrain model (DTM) or digital elevation map (DEM).

Raytracing computations for propagating HF signals require knowledge of the Earth's magnetic flux vector field at any desired location. This study uses the 11th generation model for the International Geomagnetic Reference Field (IGRF), known as IGRF-11. Its magnetic flux is modeled as the gradient of a time-varying spatial potential function. Additional information about this model can be found in International Association of Geomagnetism and Aeronomy (2019). It should be noted that an updated model for the Earth's magnetic flux vector field that is based on a 2020 epoch, IGRF-13, is now available.

3.2 | The ionosphere model

A three-parameter Chapman beta vertical profile is used to model the location-dependent electron density distribution of the ionosphere. This model regards the
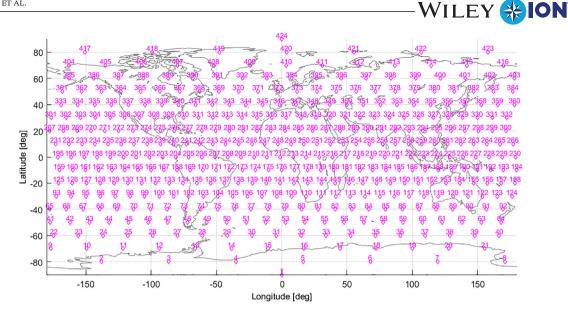
FIGURE 1 A map for the 424 nodes for an example latitude/longitude bi-quintic spline [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com and www.ion.org]

ionosphere as a medium with an altitude-dependent electron density, whose altitude density distribution is characterized by parameters that vary with latitude and longitude. Under certain assumptions for radiation, geometry, and chemistry, the Chapman profile can be regarded as the exact solution for an ionosphere density profile (Stankov et al., 2003).

For a given date and time, electron density is given by

$$N_e\left(\underline{r}\right) = \frac{VTEC\left[\phi\left(\underline{r}\right), \lambda\left(\underline{r}\right)\right]}{e \cdot h_{sf}\left[\phi\left(\underline{r}\right), \lambda\left(\underline{r}\right)\right]} \exp\left\{1 - z\left(\underline{r}\right)\right.$$
$$\left. - \exp\left[-z\left(\underline{r}\right)\right]\right\}$$
$$z\left(\underline{r}\right) = \frac{h_{alt}\left(\underline{r}\right) - h_{max}\left[\phi\left(\underline{r}\right), \lambda\left(\underline{r}\right)\right]}{h_{sf}\left[\phi\left(\underline{r}\right), \lambda\left(\underline{r}\right)\right]}, \quad (10)$$

where $\phi(\underline{r})$, $\lambda(\underline{r})$, and $h_{alt}(\underline{r})$ are, respectively, the latitude, longitude, and altitude above the WGS-84 ellipsoid of the ECEF position $\underline{r}$. $N_e(\underline{r})$ is the electron density at this position given in units of electrons/m^3. The quantity $h_{\max}[\phi(\underline{r}),\lambda(\underline{r})]$ is the altitude of the maximum electron density of the Chapman profile. The quantity $VTEC[\phi(\underline{r}),\lambda(\underline{r})]$ is the vertical total electron content – the integral of the electron density along a vertical path. The quantity $h_{sf}[\phi(\underline{r}),\lambda(\underline{r})]$ is the Chapman profile's altitude scale height. These last three functions depend on the date and the time of day in addition to latitude and longitude, but this dependence is not explicitly noted for the sake of convenience. The $N_e(\underline{r})$ function is assumed to be slowly time varying so that it can be assumed to be constant over the short interval of an HF radio wave propagation from a transmitter beacon to a user receiver.

Note that one could switch to using $N_{e,max}[\phi(\underline{r}),\lambda(\underline{r})]$ as one of the three Chapman profile parameters in place of $VTEC[\phi(\underline{r}),\lambda(\underline{r})]$. Such a switch might be appealing because the top-side contributions to $VTEC$ do not affect the bottom-side refraction that matters to the present developments. In practice, such a switch is unlikely to change the underlying performance of this concept.

The latitude/longitude dependencies of the three Chapman profile parameter functions $h_{\max}[\phi(\underline{r}),\lambda(\underline{r})]$, $VTEC[\phi(\underline{r}),\lambda(\underline{r})]$, and $h_{sf}[\phi(\underline{r}),\lambda(\underline{r})]$ are modeled using a special bi-quintic spline. It works with data that are defined at spline nodes. Spline nodes are placed at predefined latitudes and longitudes with subsets of nodes grouped into common small circles of constant latitude. Along each small circle, the spline nodes can have any desired longitude spacing, and this spacing can vary. There is no need to align node longitudes on one small circle with node longitudes on adjacent small circles. The latitude spacing between small circles is also free to vary. Figure 1 illustrates the placement of spline nodes, where each node is identified by a unique number, starting at 1 for the node that is located at the South Pole and ending at 424 for the node that is located at the North Pole. The set of spline nodes that is used in this study has been defined in a way that gives a sufficient density of nodes over North America, the region in which simulation and analysis cases are planned to be considered.

The latitude/longitude variations of the three Chapman vertical profile parameters are modeled using bi-quintic splines as described in Baumgarten (2018) and Psiaki et al. (2019). Three sets of parameters are stored for each grid node. Each set consists of the natural logarithm of a Chapman parameter's value and eight of its partial derivatives

with respect to latitude ϕ and longitude λ. Thus, a vector of $3 \times 9 = 27$ parameters, $\underline{p}_i$, is associated with the ith node as follows:

$$
\begin{aligned}
\underline{p}_i = \bigg[& \ln\left(h_{\max_i}\right), \frac{\partial \ln\left(h_{\max_i}\right)}{\partial \lambda}, \frac{\partial \ln\left(h_{\max_i}\right)}{\partial \phi}, \\
& \frac{\partial^2 \ln\left(h_{\max_i}\right)}{\partial \lambda^2}, \frac{\partial^2 \ln\left(h_{\max_i}\right)}{\partial \lambda \partial \phi}, \frac{\partial^2 \ln\left(h_{\max_i}\right)}{\partial \phi^2}, \\
& \frac{\partial^3 \ln\left(h_{\max_i}\right)}{\partial \lambda^2 \partial \phi}, \frac{\partial^3 \ln\left(h_{\max_i}\right)}{\partial \lambda \partial \phi^2}, \frac{\partial^4 \ln\left(h_{\max_i}\right)}{\partial \lambda^2 \partial \phi^2}, \\
& \ln\left(h_{sf_i}\right), \ldots, \frac{\partial^4 \ln\left(h_{sf_i}\right)}{\partial \lambda^2 \partial \phi^2}, \\
& \ln\left(VTEC_i\right), \ldots, \frac{\partial^4 \ln\left(VTEC_i\right)}{\partial \lambda^2 \partial \phi^2} \bigg]^T.
\end{aligned} \tag{11}
$$

Note that special shortened forms of this parameter vector apply at the North and South Poles. These special forms have only 18 elements rather than 27 elements.

The natural logarithms of the three Chapman profile parameters are modeled rather than the parameters themselves as an *ad hoc* means of ensuring that the final parameter values will be positive. The spline models of their natural logarithms can take on any real values. The parameters themselves will be positive after their splined natural logarithms have been input to exponential calculations.

Given the latitude ϕ_0 and the longitude λ_0 of a point at which one wants to compute electron density (and possibly various of its partial derivatives), the spline calculations use the nearest four bi-quintic spline nodes that lie northwest, northeast, southwest, and southeast of (ϕ_0, λ_0). The details of the spline evaluation procedure are given in Baumgarten (2018) and Psiaki et al. (2019).

It is important to note that the simplistic Chapman model ignores the possibility of distinct D and E layers, including a sporadic E layer. This level of simplification may produce unsatisfactory results if working with real daytime data, but it is reasonable to use a Chapman profile at this stage of simulation- and analysis-based study of the proposed system's potential accuracy.

3.3 | The ionosphere parameters variability matrix

Let $\underline{p}$ denote a stacked vector of all 424 $\underline{p}_i$ vectors. This means that a given vector $\underline{p}$ stores information that describes the entire electron density distribution near the Earth at a particular time. Understandably, $\underline{p}$ vectors that describe electron density distributions at different times take different values. While the values that the various

terms of $\underline{p}$ take may vary significantly with time, the manner at which they vary reflects the fact that those values represent physical phenomena. For instance, at a specific time, values for a particular Chapman parameter at neighboring spline nodes, that are relatively near, are expected not to differ by much, while values that are measured for one of its derivatives at a single spline node, but at different yet close times, are similarly expected to be close.

The batch-filtering algorithm that is used in this study requires a model for the likely correlations between the various terms of $\underline{p}$. This model should effectively embody the manner at which these terms vary in time and the interdependencies between them. One method for generating such a model is computing an empirical covariance matrix through the processing of parameterized models of the ionosphere that have been computed for a large series of times. The International Reference Ionosphere (IRI) model was used to compute the best-fit Chapman parameter values four times a day throughout the calendar year 2009. See Bilitza (2001) and Reinisch and Bilitza (2004) on current IRI modeling and Bilitza (2011) on further improvement efforts. See Psiaki et al. (2015) on the Chapman parameter fitting procedure. The resulting $365 \times 4 = 1,460$ parameter vectors were used to compute a covariance matrix for $\underline{p}$, i.e., for the natural logarithm and natural logarithm partial derivatives of all three Chapman model parameters at the spline nodes. This matrix is defined to be the ionospheric parameters' *variability matrix*, designated M_0 throughout this paper. It characterizes the likely variability of the $\underline{p}$ vector over a year, and it contains information about the correlations between elements of $\underline{p}$. It should be noted that M_0, which has 11,430 rows and 11,430 columns, contains much more information than is required when considering a typical HF navigation problem. The matrix M is an $N_p \times N_p$ covariance matrix that was constructed from the M_0 matrix by retaining only the rows and columns that are associated with the set of applicable spline nodes, i.e., spline nodes that define grid cells through which propagating rays travel. The value N_p is the number of ionosphere parameters that apply for a particular problem setup – typically in the order of hundreds.

4 | RAY PATHS, MULTIPLE-HOP CALCULATIONS, AND MEASUREMENT MODELS

4.1 | Definitions

In the scope of this study, it is assumed that waves are perfectly reflected from the Earth's surface in a specular
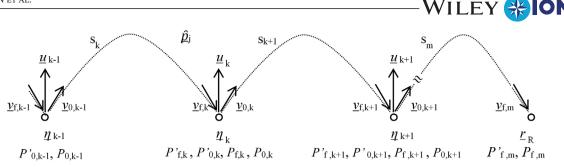
FIGURE 2 The ray-path definitions and notation

manner at *bounce points*. This simplification of the quasi-isotropic nature of ground reflections has been chosen for its relative simplicity in the multi-hop path analysis. This model will most likely need to change when dealing with real HF data to characterize ground reflections' sensitivity to surface conditions and polarization.

The position of the kth bounce point in Cartesian coordinates is denoted by $\underline{\eta}_k$. The unit vector that is perpendicular to the Earth's surface at bounce point k is called the *bounce point normal vector* and is denoted by $\underline{u}_k$. The ray-path direction from which a signal approaches bounce point k is $\underline{v}_{f,k}$. The direction of the reflected signal at bounce point k is $\underline{v}_{0,k}$. The curved signal trajectory between the transmitter and the first bounce point, between two sequential bounce points, or between the final bounce point and the receiver is termed a *ray hop*. The kth ray hop is denoted s_k. An ordered sequence of ray hops that starts at a transmitter and ends at the location of the receiver, $\underline{r}_R$, constitutes a *ray path*.

Figure 2 illustrates these definitions, showing three sequential bounce points, the receiver location, the ray hops connecting them, and other terms. The associated vector $\underline{\hat{p}}_j$ consists of all ionosphere parameters that apply in the vicinity of the jth ray path that is illustrated in the figure. All $P'_{x,y}$ and $P_{x,y}$ terms refer, respectively, to range-equivalent group delays and beat carrier phases that will be considered in a later discussion.

4.2 | Bounce-point equations

Three equations are used to implicitly define each bounce point. The kth Type-A constraint equation requires that the kth bounce point lie on the Earth. The Type-B constraint equation enforces co-planarity between the directional vector of the incoming ray-hop signal as it approaches the bounce point, the directional vector of the reflected ray hop, and the normal vector to the Earth's surface at the given bounce point. For the kth bounce point, which links the kth ray hop and the $(k+1)^{st}$ ray hop, the following equation definition for $g_{B,k}$ applies:

$$0 = g_{B,k}\left(\underline{v}_{f,k}, \underline{v}_{0,k}, \underline{\hat{\eta}}_j, \underline{u}_k\left(\underline{\hat{\eta}}_j\right)\right)$$

$$= \underline{u}_k\left(\underline{\hat{\eta}}_j\right) \cdot \left(\underline{v}_{f,k} \times \underline{v}_{0,k}\right), \tag{12}$$

where $\underline{\hat{\eta}}_j$ is a stacked vector of all $\underline{\eta}_k$ bounce-point locations of that ray path, and $\underline{u}_k$ is the outward unit vector normal to the Earth's surface at the kth bounce point (Baumgarten, 2018).

Each Type-C equation constrains the normal vector to the Earth at the bounce point to bisect the angle between the incoming and reflected ray hops. It can be written in the form:

$$0 = g_{C,k}\left(\underline{v}_{f,k}, \underline{v}_{0,k}, \underline{\hat{\eta}}_j, \underline{u}_k\left(\underline{\hat{\eta}}_j\right)\right)$$

$$= \underline{u}_k\left(\underline{\hat{\eta}}_j\right) \cdot \left(\left\|\underline{v}_{f,k}\right\|\underline{v}_{0,k} + \left\|\underline{v}_{0,k}\right\|\underline{v}_{f,k}\right). \tag{13}$$

Finally, the set of three equations that defines the kth bounce point of a given ray path can be written in the following shorthand form:

$$\underline{0} = \underline{g}_k = \underline{g}_k\left(\underline{v}_{f,k}, \underline{v}_{0,k}, \underline{\eta}_k, \underline{u}_k\left(\underline{\hat{\eta}}_j\right)\right). \tag{14}$$

Recognizing that a signal's trajectory within a single ray hop, and in particular its directional vectors $\underline{v}_0$ and $\underline{v}_f$, depends on the location of the hop's start and end points and on the values taken by the ionosphere parameters that apply in the vicinity of that hop, Equation (14) can be rewritten as

$$\underline{0} = \underline{\tilde{g}}_k\left(\underline{\hat{\eta}}_j, \underline{r}_R, \underline{\hat{p}}_j\right), \tag{15}$$

where the formulation in Equation (15) implicitly uses the ray-tracing calculations for each of the single hops in

order to compute the corresponding $\underline{v}_{0,k-1}$ and $\underline{v}_{f,k}$ directions from the given bounce-point locations defined from components of $\widehat{\underline{\eta}}_j$ and the transmitter location for the initial hop or the receiver location $\underline{r}_R$ for the final hop. Both formulations of Equations (14) and (15) for the set of three equations are used in this study. The first is used with most batch-filtering Gauss-Newton process-related calculations and the second with the ray-path solver that is described in this section. Additional details are available in Baumgarten (2018).

4.3 | Single-hop calculations

Given the signal trajectory's known start and end locations for a single hop, and given a set of applicable ionosphere parameters, one can determine the ray-hop trajectory by determining the initial state $\underline{X}_0$ of the raytracing differential equation in Equation (8) that applies at the beginning of the hop's trajectory and the total signal range-equivalent group delay P'_f for which the signal ultimately arrives at the known end location. This two-point boundary value problem (TPBVP) and an iterative solution method are thoroughly discussed in Psiaki (2019).

Baumgarten (2018) puts the discussion in Psiaki (2019) in the context of the current study. It shows how the problem is solved based on the principle of a zero-valued Hamiltonian, which is kept fixed throughout the signal propagation along its trajectory. It additionally describes how the problem is solved using a Newton method and outlines the derivation of the sensitivity matrices that are required for the computation of each Newton step within this process. The Newton's method solution uses ray-path sensitivity calculation of partial derivatives with respect to the initial unknown wave vector and with respect to the unknown total range-equivalent group delay. Related calculations can be used to compute the partial derivative sensitivity of the resulting ray path with respect to changes in the initial or final bounce points and with respect to changes in the ionosphere parameters that affect the single hop. These latter sensitivities are not needed in order to solve a given single ray hop, but they are needed by the solution procedure for a multi-hop ray path and by the partial derivative sensitivity calculations of the batch filter that uses ray-path solutions to model HF pseudorange and beat carrier-phase observables.

4.4 | Multiple-hop calculations

The work that is presented in this subsection utilizes the single-hop method of Psiaki (2019) as reviewed in the previous subsection. In the following discussion, single-hop calculations are extended to multiple-hop calculations that determine the bounce points' locations, the group delays, the range-equivalent beat carrier phases, and these quantities' partial derivative sensitivities to inputs.

Determination of the ray path for an HF signal that is traversing from a transmitter beacon to a receiver involves the solution of coupled, nonlinear equations that define the physical characteristics of its trajectory. Given the locations of the receiver and the transmitter, the number of ray-path hops connecting them, and a model for the ionosphere, the objective is to solve for the position of all of the ground bounce points in $\widehat{\underline{\eta}}_j$ in a way that satisfies the governing reflection equations while also solving for the set of single hops that properly connect the bounce points. Ultimately, bounce-points solutions in $\widehat{\underline{\eta}}_j$ depend on the receiver position $\underline{r}_R$ and the relevant ionosphere parameters $\hat{p}_j$.

An algorithm for determining this nonlinear function $\widehat{\underline{\eta}}_j(\underline{r}_R, \hat{p}_j)$ has been developed based on the implicit equations that define it. It is called a *ray-path solver*. The ray-path solver assumes fixed known locations for the signal's start and end points, fixed ionosphere parameters, and a known number of sequenced ray hops that constitute the ray path. The ray-path solver's standard outputs are the locations of the bounce points. Auxiliary outputs include the ray-traced single hops' trajectories between each pair of bounce points along with the pseudorange and beat carrier-phase increments along each single hop. If required by the batch filter, associated computations can determine the partial derivatives of bounce-point locations, total group delays, and beat carrier-phase increments with respect to the location of the ray path's end point location $\underline{r}_R$ and with respect to the ionosphere parameters in $\hat{p}_j$

The multi-hop solution is obtained using Newton's iterative method to solve

$$\underline{0} = \widehat{\underline{\tilde{g}}}_j\left(\widehat{\underline{\eta}}_j, \underline{r}_R, \widehat{\underline{p}}_j\right) \tag{16}$$

through minimization of

$$J_{RP}\left(\widehat{\underline{\eta}}_j\right) = \frac{1}{2}\widehat{\underline{\tilde{g}}}_j\left(\widehat{\underline{\eta}}_j, \underline{r}_R, \widehat{\underline{p}}_j\right)^T \widehat{\underline{\tilde{g}}}_j\left(\widehat{\underline{\eta}}_j, \underline{r}_R, \widehat{\underline{p}}_j\right) \tag{17}$$

with respect to $\widehat{\underline{\eta}}_j$. Here, $\widehat{\underline{\tilde{g}}}_j$ denotes the vector of m stacked Equation (15)-type formulations of the set of three reflection equations that apply at each of the m bounce points of ray path j. The ray path consists of $m+1$ connected ray hops. The stacked vector $\widehat{\underline{\eta}}_j$ contains the $3m$ coordinates of the m ray paths' bounce points, and the vector function $\widehat{\underline{\tilde{g}}}_j$ has $3m$ elements. Therefore, Equation (16) is a system of $3m$ nonlinear equations in $3m$ unknowns.

The iterative solution procedure uses linearization about a current guess to compute a solution increment. It takes a step along the resulting search direction with a step-length scaling in the $\widehat{\underline{\eta}}_j$ space that is chosen to ensure that the Equation (17) new cost at the new guess of the solution is lower than the cost at the previous guess. A line search starts by determining the Newton step through linearization of Equation (16) with respect to $\widehat{\underline{\eta}}_j$, where all terms are evaluated at a current guess $\widehat{\underline{\eta}}_{j,guess}$

$$\left. \frac{D\widehat{\tilde{\underline{g}}}_j}{D\widehat{\underline{\eta}}_j} \right|_{\widehat{\underline{\eta}}_{j,guess}} \Delta \widehat{\underline{\eta}}_j = -\widehat{\tilde{\underline{g}}}_j \left(\widehat{\underline{\eta}}_{j,guess}, \underline{r}_R, \widehat{\underline{p}}_j \right), \qquad (18)$$

and where D denotes the total derivative operator. The solution for $\Delta\widehat{\underline{\eta}}_{j,guess}$, which constitutes the Gauss-Newton step (or correction vector), is computed by matrix inversion. For bounce point k, the required set of sensitivity matrices that are included in the leftmost term of Equation (18) is obtained through computation of the total derivative

$$\frac{D\tilde{\underline{g}}_k}{D\underline{\eta}_l} = \frac{\partial \tilde{\underline{g}}_k}{\partial \underline{v}_{f,k}} \frac{\partial \underline{v}_{f,k}}{\partial \underline{\eta}_l} + \frac{\partial \tilde{\underline{g}}_k}{\partial \underline{v}_{0,k}} \frac{\partial \underline{v}_{0,k}}{\partial \underline{\eta}_l} + \frac{\partial \tilde{\underline{g}}_k}{\partial \underline{\eta}_l}, \qquad (19)$$

where $\tilde{\underline{g}}_k$ is the subset of the elements of $\widehat{\tilde{\underline{g}}}_j$, consisting of the three equations that apply at bounce point η_k. η_l is the lth bounce point of that ray path where l takes the values k-1, k, and k+1. Computations of $\partial\tilde{\underline{g}}_k/\partial\underline{v}_{0,k}$, $\partial\tilde{\underline{g}}_k/\partial\underline{v}_{f,k}$, and $\partial\tilde{\underline{g}}_k/\partial\underline{\eta}_k$ are analytical and therefore immediate as noted earlier. However, computations of $\partial\underline{v}_{f,k}/\partial\underline{\eta}_{k-1}$, $\partial\underline{v}_{f,k}/\partial\underline{\eta}_k$, $\partial\underline{v}_{0,k}/\partial\underline{\eta}_k$, and $\partial\underline{v}_{0,k}/\partial\underline{\eta}_{k+1}$ are implemented as auxiliaries of numerical raytracing as described in Baumgarten (2018) and Psiaki (2019).

An initial guess for $\widehat{\underline{\eta}}_j$ is generated using several methods that are described in Baumgarten (2018). The possible methods include the use of the simplified ray-path solver of Baumgarten and Psiaki (2017), the use of a latitude/longitude/altitude-dependent thin-shell ionosphere model, and the use of a constant-altitude thin-shell ionosphere model.

4.5 | Ray paths' feasibility and solution uniqueness

Every ray path is evaluated for physical feasibility. Physical feasibility concerns the question of whether there exists a ray path between the given transmitter and receiver locations with the given number of hops at the given carrier frequency and for a given set of ionospheric parameters. For simulated test cases, feasibility is evaluated by trying to compute a raytracing solution using the true ionosphere model. The answer for the feasibility question is often not straightforward as it consists of the questions of whether a solution for a given set of inputs exists and, if so, whether it can be found with the ray-path solver. In the absence of an ability to distinguish between a negative answer for the two questions, a failure in obtaining a solution for $\widehat{\underline{\eta}}_j$ during this phase of assessment is generally regarded as an indication that the path is physically infeasible for the given inputs.

Ray-path solution uniqueness is a second matter of concern. Multiple solutions are theoretically possible if the cost function has multiple minima that are zero as demonstrated in Australian Government (2016). In the early work that utilized a simplified ray-path model (Baumgarten & Psiaki, 2017), a given set of inputs that included transmitter and receiver position, an ionosphere model, and the number of ray hops sometimes yielded more than one possible solution. Such observations have not been made so far with the full, raytracing-based model of the present paper. The current study, therefore, does not consider the possibility of having more than one solution for $\widehat{\underline{\eta}}_j$.

4.6 | Measurement models

In the context of this study, range measurements are based on signal propagation time measurements and wave-phase measurements. Errors in processing the measured data may arise from 1) errors in the modeled signal propagation paths, 2) clock synchronization errors, and 3) signal-processing related errors. The first two types of error sources are addressed through proper modeling of these error sources in the batch-filtering-based algorithm. The third error source is generalized as a noise term in the following measurement equations.

A typical signal runs from the transmitter, traverses through the ionosphere in a refraction-based curved trajectory, bounces off of Earth, and eventually arrives at the receiver. Let $\rho_{g,j} = P'_{f,m(j)}$ be the true total range-equivalent group delay of the jth ray path, which equals the true signal propagation time multiplied by the speed of light c. Let $y_{g,j}$ be the measured range-equivalent group delay of that ray path, which equals the speed of light multiplied by the difference between the measured reception time according to the erroneous receiver clock and the true transmission time according to a calibrated beacon

transmitter clock. Let δ be the receiver clock's offset and let $\underline{x}_g = [\underline{r}_R{}^T, c\delta]^T$ denote the vector of unknown receiver position components and the range-equivalent clock offset. Then, the jth group delay measurement equation can be written as

$$
\begin{aligned}
y_{g,j} &= h_{g,j}\left(\underline{x}_g, \widehat{\underline{p}}_j\right) + \nu_{g,j} = \check{h}_{g,j}\left(\underline{r}_R, \widehat{\underline{p}}_j\right) + c\delta + \nu_{g,j} \\
&= \tilde{h}_{g,j}\left[\underline{r}_R, \widehat{\underline{\eta}}_j\left(\underline{r}_R, \widehat{\underline{p}}_j\right), \widehat{\underline{p}}_j\right] + c\delta + \nu_{g,j},
\end{aligned} \tag{20}
$$

where the computed functions $\check{h}_{g,j}$ and $\tilde{h}_{g,j}$ both model the true range-equivalent group delay of the jth ray path. $\nu_{g,j}$ is a zero-mean measurement noise term that embodies the effect of signal-processing related errors.

The measurement model in Equation (20) applies for a total of N measured pseudoranges in a given navigation/ionosphere-correction problem. For convenience in batch estimation, this model is stacked into an N-dimensional vector equation model of all the measurements. Let $\widehat{\underline{p}}$ equal the union of all $\widehat{\underline{p}}_j$ vectors applying for all N ray paths. The stacked measurement model vector equation takes the form:

$$
\underline{y}_g = \underline{h}_g\left(\underline{x}_g, \widehat{\underline{p}}\right) + \underline{\nu}_g; \quad E\left[\underline{\nu}_g\right] = \underline{0}; \quad E\left[\underline{\nu}_g \underline{\nu}_g{}^T\right] = R_g \tag{21}
$$

with measurement error covariance matrix R_g typically a diagonal matrix. Note that it is possible for two or more of the N ray paths to originate from the same transmitter location. In that case, either the signal transmission frequency, the number of hops, or both must be different in order for the measurements to provide independent information.

The second type of measurement used in this study, beat carrier phase, is based on a comparison between measured changes in the received signal's phase and changes in the phase of a receiver-generated nominal replica signal. In effect, the beat carrier phase is the negative of the time integral of the received carrier Doppler shift (Bennett, 1967). This measurement involves an unknown bias term that originates from its integral nature, i.e., it is an unknown integration constant. Let $\rho_{c,j} = P_{f,m(j)}$ be the total true range-equivalent beat carrier phase of the jth ray path, and let $y_{c,j}$ be the measured range-equivalent beat carrier phase of that ray path. Recall that P is computed by integrating the differential equation in Equation (4). Let $\lambda_{w,j}$ be the corresponding signal's wavelength, and let $\beta_{i(j)}$ be an unknown bias term in units of carrier cycles. Then, the jth beat carrier-phase measurement equation can be written as

$$
\begin{aligned}
y_{c,j} &= h_{c,j}\left(\underline{x}_c, \widehat{\underline{p}}_j\right) + \nu_{c,j} = \check{h}_{c,j}\left(\underline{r}_R, \widehat{\underline{p}}_j\right) + c\delta \\
&\quad + \lambda_{w,j}\beta_{i(j)} + \nu_{c,j} = \tilde{h}_{c,j}\left[\underline{r}_R, \widehat{\underline{\eta}}_j\left(\underline{r}_R, \widehat{\underline{p}}_j\right), \widehat{\underline{p}}_j\right] + c\delta \\
&\quad + \lambda_{w,j}\beta_{i(j)} + \nu_{c,j},
\end{aligned} \tag{22}
$$

where $\underline{x}_c = [\underline{r}^T, c\delta, \underline{\beta}^T]^T$ and where the computed functions $\check{h}_{c,j}$ and $\tilde{h}_{c,j}$ both model the ionosphere-refraction-induced range-equivalent carrier-phase change of the jth ray path, $\rho_{c,j}$. The vector $\underline{\beta}$ consists of all unknown bias terms that apply for all N beat carrier-phase measurements.

The integer function $i(j)$ in the index of β maps ray-path indices j to indices of their corresponding terms in $\underline{\beta}$. The ability to map a common bias term to multiple beat carrier-phase measurements is needed because the beat carrier-phase data are not useful unless multiple measurements are made with a common bias. This can be accomplished by transmitting a signal with a frequency time history that follows a smoothed staircase pattern with a known continuous phase time history at the transmitter. Coherent reception of this signal with a PLL that tracks phase followed by differencing of the measured phase time history from the known transmitted phase time history results in a set of beat carrier-phase measurements that have a common bias in terms of cycles, a common transmitter location, and a common number of hops, but a different transmission frequency. If the beat carrier-phase measurement data are used from a set of two or more times when the signal is temporarily staying at a constant frequency, but with a different constant frequency for each of those times, then the model in Equation (22) applies for more than one value of j that map to an identical bias index $i(j)$. It is assumed that a given transmitter's smoothed stair-stepping frequency transmission time history will occur over a relatively short time window, perhaps just 10 msec, so that the receiver location, the receiver clock offset, and the ionosphere model can be assumed to remain constant during that window.

As with group delay measurements, carrier-phase measurement equations are stacked into an N-dimensional vector measurement model equation that takes the form:

$$
\underline{y}_c = \tilde{\underline{h}}_c\left(\underline{r}_R, \widehat{\underline{p}}\right) + \underline{1}c\delta + \Lambda\underline{\beta} + \underline{\nu}_c
$$

$$
E\left[\underline{\nu}_c\right] = \underline{0}; \; E\left[\underline{\nu}_c\underline{\nu}_c{}^T\right] = R_c, \tag{23}
$$

where Λ is an $N \times N_\beta$ matrix and N_β is the dimension of $\underline{\beta}$. The random measurement noise vector, $\underline{\nu}_c$, is characterized by its mean (assumed to be zero) and its covariance matrix R_c.

Finally, both types of measurement vector, the first for the range-equivalent group delays and the second for the range-equivalent beat carrier phases, can be stacked into a single $2N$-dimensional measurement vector. The same can be done with the vector functions $\underline{h}_g$ and $\underline{h}_c$ and with the noise vectors $\underline{\nu}_g$ and $\underline{\nu}_c$:

$$\underline{y} = \begin{bmatrix} \underline{y}_g \\ \underline{y}_c \end{bmatrix}; \quad \underline{h}\left(\underline{x}_c, \widehat{\underline{p}}\right) = \begin{bmatrix} \underline{h}_g\left(\underline{x}_g, \widehat{\underline{p}}\right) \\ \underline{h}_c\left(\underline{x}_c, \widehat{\underline{p}}\right) \end{bmatrix}; \quad \underline{\nu}_z = \begin{bmatrix} \underline{\nu}_g \\ \underline{\nu}_c \end{bmatrix};$$

$$R = E\left[\underline{\nu}_z \underline{\nu}_z^T\right] = \begin{bmatrix} R_g & 0 \\ 0 & R_c \end{bmatrix}. \tag{24}$$

Note that $\underline{x}_g \subset \underline{x}_c$ so that $\underline{h}$ is conveniently defined as a function of $\underline{x}_c$. The resulting measurement model takes the form:

$$\underline{y} = \underline{h}\left(\underline{x}_c, \widehat{\underline{p}}\right) + \underline{\nu}_z. \tag{25}$$

4.7 | Measurement model sensitivity matrices

Gradient-based nonlinear estimation algorithms, such as batch least-squares, require partial derivatives of the measurement model with respect to the unknown estimated quantities. These sensitivities must be computed at a succession of improved guesses of the optimal estimates of the unknowns. In the present context, the required partial derivatives are those of each h_j measurement model function with respect to the elements of the unknown $\underline{x}$ and $\hat{p}$ vectors. Derivatives with respect to the elements of $\underline{r}$ and $\hat{p}$ require special calculations. The sensitivity of the jth range measurement to the input variables $\underline{r}_R$ and $\hat{\underline{p}}_j$ is

$$\frac{D\underline{h}_j}{D\underline{r}_R} = \frac{\partial \tilde{\underline{h}}_j}{\partial \widehat{\underline{\eta}}_j} \frac{\partial \widehat{\underline{\eta}}_j}{\partial r_R} + \frac{\partial \tilde{\underline{h}}_j}{\partial \underline{r}_R}$$

$$\frac{D\underline{h}_j}{D\widehat{\underline{p}}_j} = \frac{\partial \tilde{\underline{h}}_j}{\partial \widehat{\underline{\eta}}_j} \frac{\partial \widehat{\underline{\eta}}_j}{\partial \widehat{\underline{p}}_j} + \frac{\partial \tilde{\underline{h}}_j}{\partial \widehat{\underline{p}}_j}, \tag{26}$$

where

$$\frac{\partial \widehat{\underline{\eta}}_j}{\partial \underline{r}_R} = -\left(\frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{v}}_{f,j}} \frac{\partial \widehat{\underline{v}}_{f,j}}{\partial \widehat{\underline{\eta}}_j} + \frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{v}}_{0,j}} \frac{\partial \widehat{\underline{v}}_{0,j}}{\partial \widehat{\underline{\eta}}_j} + \frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{u}}_j} \frac{\partial \widehat{\underline{u}}_j}{\partial \widehat{\underline{\eta}}_j} + \frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{\eta}}_j}\right)^{-1}$$

$$\times \left(\frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{v}}_{0,j}} \frac{\partial \widehat{\underline{v}}_{0,j}}{\partial \underline{r}_R}\right)$$

$$\frac{\partial \widehat{\underline{\eta}}_j}{\partial \widehat{\underline{p}}_j} = -\left(\frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{v}}_{f,j}} \frac{\partial \widehat{\underline{v}}_{f,j}}{\partial \widehat{\underline{\eta}}_j} + \frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{v}}_{0,j}} \frac{\partial \widehat{\underline{v}}_{0,j}}{\partial \widehat{\underline{\eta}}_j} + \frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{u}}_j} \frac{\partial \widehat{\underline{u}}_j}{\partial \widehat{\underline{\eta}}_j} + \frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{\eta}}_j}\right)^{-1}$$

$$\times \left(\frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{v}}_{f,j}} \frac{\partial \widehat{\underline{v}}_{f,j}}{\partial \widehat{\underline{p}}_j} + \frac{\partial \widehat{\underline{g}}_j}{\partial \widehat{\underline{v}}_{0,j}} \frac{\partial \widehat{\underline{v}}_{0,j}}{\partial \widehat{\underline{p}}_j}\right). \tag{27}$$

The column vector $\hat{g}_j$ is a stacked vector consisting of the m three-term g_k function vectors associated with the m bounce points of ray path j. Similarly, $\hat{v}_{0,j}$, $\hat{v}_{f,j}$, and $\hat{u}_j$ are stacked column vectors for the m three-term vectors of the approaching signals, reflected signals, and bounce-point normal vectors, respectively.

Of the terms on the right-hand side of Equation (27), the following terms can be evaluated analytically through differentiation of the bounce-point equations that were introduced earlier: $\partial \hat{g}_j / \partial \widehat{\underline{\eta}}_j$, $\partial \hat{g}_j / \partial \hat{v}_{0,j}$, $\partial \hat{g}_j / \partial \hat{v}_{f,j}$, and $\partial \hat{g}_j / \partial \hat{u}_j$. Other terms, however, can only be evaluated in tandem with the raytracing calculations that determine single ray hops: $\partial \hat{v}_{f,j} / \partial \hat{p}_j$, $\partial \hat{v}_{0,j} / \partial \hat{p}_j$, $\partial \hat{v}_{f,j} / \partial \widehat{\underline{\eta}}_j$, $\partial \hat{v}_{0,j} / \partial \widehat{\underline{\eta}}_j$, and $\partial \hat{v}_{0,j} / \partial r_R$. The computation of these terms has been described in Baumgarten (2018) and Psiaki (2019).

5 | BATCH ESTIMATION OF RECEIVER POSITION, RECEIVER CLOCK OFFSET, AND IONOSPHERE PARAMETERS

A batch filter has been developed. It estimates $\underline{x}_c$ and $\underline{p}$ by minimizing a cost function that includes weighted squared differences between the measurements and their modeled values and between the estimated $\underline{p}$ elements and their *a priori* estimates.

5.1 | Batch-filter problem definition

In the general case, the batch-filtering problem seeks the values that jointly minimize the cost function:

$$J_1\left(\underline{x}_c, \underline{p}\right) = \frac{1}{2}\left[\underline{y} - \underline{h}\left(\underline{x}_c, \underline{p}\right)\right]^T R^{-1}\left[\underline{y} - \underline{h}\left(\underline{x}_c, \underline{p}\right)\right]$$

$$+ \frac{1}{2}\zeta(\underline{p} - \underline{\bar{p}})^T M^{-1}\left(\underline{p} - \underline{\bar{p}}\right), \tag{28}$$

where $\underline{y}$ is the $2N$x1 stacked vector of the N measured pseudoranges and N measured range-equivalent beat carrier phases for the given N ray paths. R is the square, symmetric, $2N$-by-$2N$, positive-definite measurement error covariance matrix. $\underline{\bar{p}}$ is the *a priori* estimate of the ionosphere

parameter vector, and M is the square, symmetric, positive definite covariance matrix that models the uncertainty in the *a priori* ionosphere parameter vector $\bar{p}$. The elements of p consist of ionosphere parameters, which apply in the vicinity of the unknown, true signal ray paths. ζ is a positive scaling parameter that effectively re-scales the inverse of the *a priori* ionosphere parameter error covariance matrix. Its role is described in more detail in Baumgarten (2018).

The batch least-squares cost function of Equation (28) does not include *a priori* values of the elements of x_c with penalties for differences between those values and the estimated x_c. This means that no prior knowledge about receiver position, receiver clock offset, or beat carrier-phase biases is assumed.

The minimizing solution to this estimation problem is equivalent to the optimal least-squares solution to the following over-determined system of nonlinear equations:

$$\begin{bmatrix} R^{-1/2}y \\ \sqrt{\zeta}M^{-1/2}\bar{p} \end{bmatrix} = \begin{bmatrix} R^{-1/2}h\left(x_c,p\right) \\ \sqrt{\zeta}M^{-1/2}p \end{bmatrix} + \nu_1, \quad (29)$$

where $R^{-1/2}$ and $M^{-1/2}$ are the inverses of the Cholesky-factor square roots of, respectively, the matrices R and M, and where ν_1 is a zero-mean, identity-covariance Gaussian random error vector whose norm squared is minimized by the batch solution. Baumgarten (2018) provides additional details about this nonlinear least-squares problem.

In some cases, it is desirable to solve for the unknown ionosphere model (and, potentially, for the unknown beat carrier-phase biases) while the receiver location and clock offset are assumed known and fixed (or, at least, closely monitored and corrected). In such cases, the optimization problem takes the general form:

$$J_4(\beta, p) = \frac{1}{2}\left[y - h(\beta, p)\right]^T R^{-1}\left[y - h(\beta, p)\right]$$
$$+ \frac{1}{2}\zeta(p - \bar{p})^T M^{-1}(p - \bar{p}). \quad (30)$$

If beat carrier-phase measurements are not processed, then the problem reduces to estimation of ionosphere parameters only.

5.2 | A modified Gauss-Newton solution algorithm

The Gauss-Newton method has been used to solve this estimation problem by finding the minimum of the cost function in Equation (28). This method is described in Gill et al. (1995) and Nocedal and Wright (2006). It is additionally discussed in the context of convex function optimization through nonlinear programing in Bertsekas (1997). Adap-

tations to this method have been made in order to address some special characteristics of the present problem.

Each iteration starts with guesses of the optimizing values of x_c and p. The Gauss-Newton method linearizes Equation (29) about these guessed values. Next, it solves the resulting over-determined linear least-squares problem to get candidates for improved solution guesses of x_c and p. Finally, it searches along the line in $[x_c;p]$ space from the old guess to the candidate new guess in order to find a new guess that reduces the cost $J_1(x_c,p)$.

Linearization of Equation (29) about the current guess for the unknown x_c and p, $x_{c,guess}$ and p_{guess}, takes the form

$$\begin{bmatrix} R^{-\frac{1}{2}} & 0 \\ 0 & \sqrt{\zeta}M^{-1/2} \end{bmatrix} \begin{bmatrix} \frac{\partial h\left(x_c,p\right)}{\partial x_c} & \frac{\partial h\left(x_c,p\right)}{\partial p} \\ 0 & I \end{bmatrix}\Bigg|_{x_{c,guess},p_{guess}} \begin{bmatrix} \Delta x_c \\ \Delta p \end{bmatrix}$$

$$= \begin{bmatrix} R^{-\frac{1}{2}} & 0 \\ 0 & \sqrt{\zeta}M^{-1/2} \end{bmatrix} \begin{bmatrix} y - h(x_{c,guess}, p_{guess}) \\ \bar{p} - p \end{bmatrix}. \quad (31)$$

This over-determined system of equations is solved through a short series of operations. The details of these procedures are presented in Baumgarten (2018). The sequence of procedures is repeated iteratively until the cost function is minimized.

The method used in this study deviates from the classic Gauss-Newton method in two respects. First, it uses an approach that allows the sets of considered measurements to change during the iterative process. This feature, which requires modifications to the way cost function reduction is approached, has been developed in order to deal with occasional failures in ray-path solving attempts due to a poor estimate for the location of the receiver and the ionospheric parameters, to difficulties in the numerical raytracing computation for one or more of the ray path's ray hops, or to the physical non-feasibility of the ray path. Regardless of the cause, the particular measurements that failed to be computable in the filter's model are temporarily excluded from the set of measurements that are considered. A second feature that is used with the batch filer is a measurement rejection mechanism. These are common practice in sensor-based systems due to the potential of significant, un-modeled measurement errors that affect sensor readings. In the context of this study, excessive measurement errors are handled with likelihood tests that are designed to detect and reject outliers as bad data.

Recognizing the challenges posed to the first-order Gauss-Newton method when it starts with a guess that is far from the receiver's true location, the algorithm distinguishes between two cases. In the nominal case that has been described above, the current position guess is assumed to be close to the solution. In this case, the

algorithm will consider variations in the three components of the ECEF representation of the receiver's location $\underline{r}_R$, variations in the range-equivalent receiver clock offset $c\delta$, variations in the carrier-phase measurement biases β, and variations in the ionosphere parameters at all bi-quintic spline nodes that affect the ray paths. If the current position guess is suspected of being far from the final solution, however, then only group-delay measurements will be processed. In addition, the algorithm will only consider variations in the receiver position's latitude and longitude and in its clock bias. Variations of altitude and of ionospheric model parameters are excluded in the calculation of the Gauss-Newton step. Additional details about this far-from-the-solution case are described in Baumgarten (2018).

6 | PRELIMINARY RESULTS

Assessment of the proposed navigation system's effectiveness evaluates the batch filter's performance in terms of positioning accuracy and its ability to estimate corrections to an erroneous *a priori* model of the ionosphere. A limited assessment has been performed through analyzing several test cases that differ in the number of ground stations and their placement, the number of ray paths, and the difference between true and *a priori* ionospheric models.

6.1 | Truth-model simulation

A truth-model simulation has been developed for algorithm validation and assessment and for solution accuracy analysis. The simulation enables testing of different combinations of ground station arrays, ionosphere error models, and other parameters. Computation for an $N_e(\underline{r})$ electron density truth model utilizes a Chapman profile that is fit to an IRI model for a particular time (Psiaki et al., 2015). A similar procedure is used to generate an *a priori* estimate of the ionosphere parameter vector for use in the batch filter's cost function. The values that characterize this estimate may deviate substantially from the values that characterize the truth-model ionosphere parameters. This would be the case of an inaccurate *a priori* ionosphere model.

The simulation uses truth values of the $\underline{x}$ and $\underline{p}$ vectors in the vector pseudorange and beat carrier-phase measurement models of Equations (21) and (23) to generate measurement values that are input directly into the main batch-filtering algorithm.

6.2 | Solution convergence

A key event in the execution of the main solver algorithm is identifying solution convergence. This is performed using a step magnitude criterion that is applied to the correction vector generated at every iteration of the Gauss-Newton solver. In theory, the Gauss-Newton method with line search is guaranteed to converge to a local minimum, but the minimum is not guaranteed to be global. Testing experience indicates that convergence to a local minimum that is different from the global minimum occurs rarely, if ever. Therefore, for all simulation-generated test cases, the algorithm seems to be insensitive to the initial guess and the corresponding magnitude of the initial error and is nearly guaranteed to converge to its global minimum. Validity of the latter statement is further explored and demonstrated in Baumgarten (2018).

6.3 | A posteriori position and ionosphere model accuracy

A test case that considers a setup with 32 varying-frequency signals transmitted from 11 ground stations has been studied. In this test case, parameterized ionosphere models for the true ionosphere and the *a priori* ionosphere that is input to the batch filter are similar. This is the case where a relatively accurate model for electron density in the vicinity of the transmitting stations and the receiver is available, possibly due to the use of one of the ionosphere-characterization methods mentioned in the introduction. Position accuracy for this test case is within 30 meters horizontal and 2 meters vertical 90% of the time, where the mean error is within meters from the receiver's true location. These results imply that, with a sufficient number of signals received and dual group-delay/beat-carrier-phase measurement processing, the achieved accuracy is adequate for the purpose of navigation and guidance for many significant applications.

A second test case, that considers only 17 transmitted signals, exhibited somewhat inferior position accuracy with errors that are roughly three times larger horizontally. Vertical accuracy, however, remained in the order of meters. In a third test case, that is characterized by a less accurate *a priori* model for the ionosphere, the mean error rose from several meters to 60 meters, resulting in total errors up to a hundred meters.

With respect to the filter's ability to apply corrections to an erroneous *a priori* ionosphere model, the algorithm has proven successful in reducing errors in the ionosphere model parameters. As one might expect, smaller errors have been observed near the locations where ray paths travel through the ionosphere. *A posteriori* ionosphere models are improved for all test cases in comparison to their *a priori* counterparts, evident in smaller *a posteriori* errors computed for the three Chapman profile parameters in the vicinity of the true ray paths.

Additional details on the studied test cases are available in Baumgarten (2018) and Baumgarten and Psiaki (2019).

6.4 | Further performance assessment

A thorough study of performance for the developed batch filter is reported in Baumgarten (2018) and Baumgarten and Psiaki (2019) and may be published in a future journal article. It includes an analysis of a series of test cases that differ in the sets of parameters that define them. These parameters include the following: type of available measurements, number of ground stations and their placement, number of ray paths, ray-path geometry, the number of hops for each ray path, signal frequencies, true vs. *a priori* ionospheric models, receiver clock error, and the true location of the receiver. Analyses of results for these test cases explore the positioning sensitivity to scenario parameters. They also explore the extent to which errors in an *a priori* parameterized ionosphere model can be reduced. Additional analyses study other aspects of the filter's performance. These include filter convergence characteristics, scenario setup feasibility, and algorithm robustness.

7 | SUMMARY AND CONCLUSION

An algorithm has been developed that utilizes group-delay/pseudorange and beat carrier-phase measurements from HF signals propagating in the ionosphere to solve a combined positioning/ionosphere-corrections problem. These HF signals are transmitted from stationary ground-based beacons at known locations. They propagate to an over-the-horizon user receiver at an unknown location via refraction-induced bounces off of the ionosphere and, possibly, intervening reflections off of the Earth's surface.

A navigation filter estimates user position, user clock error, beat carrier-phase measurement biases, and corrections to parameters that characterize the ionosphere's three-dimensional electron density profile. The nonlinear batch least-squares estimation problem is solved using a modified Gauss-Newton method. This method has a high rate of achieving successful convergence to the optimal value of the underlying cost function.

This paper presents the main assets that have been developed in this study: physical models for the Earth and the ionosphere, a model for signals propagating in the ionosphere, two different signal measurement models, and a batch filter that estimates the user location and corrections to an *a priori* parameterized model of the ionosphere.

A limited investigation of system performance has been carried out using a truth-model simulation. Simulated test cases that consider different combinations of problem characteristics have been studied. Results indicate feasibility for the combined HF navigation/ionosphere-correction concept. It has been shown that, with sufficient availability of received signals, navigation grade accuracy for positioning may be achievable. Furthermore, *a posteriori* ionosphere model estimates are consistently improved for these cases in comparison to their *a priori* counterparts.

ORCID

Yoav Baumgarten ⬤ https://orcid.org/0000-0001-7412-0329

Mark L. Psiaki ⬤ https://orcid.org/0000-0002-3070-1217

REFERENCES

Baumgarten, Y. (2018). *Navigation using high-frequency ground beacons and ionosphere model corrections* (Doctoral dissertation). Cornell University, Ithaca.

Baumgarten, Y., & Psiaki, M. L. (2017). Navigation using high-frequency ground beacons and ionosphere model corrections. *NAVIGATION*, *64*(2), 255–271. https://doi.org/10.1002/navi.196

Baumgarten, Y., & Psiaki, M. L. (2019, March). Navigation and ionosphere characterization using high-frequency signals – Models, solution concepts, and a performance analysis. *Israel Annual Conference on Aerospace Sciences*, Haifa, Israel.

Bennett, J. A. (1967). The ray theory of doppler frequency shift. *Australian Journal of Physics*, *21*(3), 259-272. https://doi.org/10.1071/PH680259

Bennett, J. A., Dyson, P. L., & Norman, R. J. (2004). Progress in radio ray tracing in the ionosphere. *The Radio Science Bulletin*, *310*, 81–91. https://doi.org/10.23919/URSIRSB.2004.7909439

Bertsekas, D. P. (1997). Nonlinear programming. *Journal of the Operational Research Society*, *48*(3). https://doi.org/10.1057/palgrave.jors.2600425

Bilitza, D. (2001). International reference ionosphere. *Radio Science*, *36*(2), 261–275. https://doi.org/10.1029/2000RS002432

Bilitza, D. (2011). The international reference ionosphere today and in the future. *Journal of Geodesy*, *85*(12), 909–920. https://doi.org/10.1007/s00190-010-0427-x

Cummer, S. A. (2000). Modeling electromagnetic propagation in the Earth-ionosphere waveguide. *IEEE Transactions on Antennas and Propagation*, *48*(9), 1420–1429. https://doi.org/10.1109/8.898776

Fridman, S., Nickisch, L., Aiello, M., & Hausman, M. (2006). Real-time reconstruction of the three-dimensional ionosphere using data from a network of GPS receivers. *Radio Science*, *41*(5). https://doi.org/10.1029/2005RS003341

Gill, P. E., Murray, W., & Wright, M. H. (1995). *Practical optimization*. Academic Press Limited.

Haselgrove, J. (1954). *Report of conference on the physics of the ionosphere*. London Physical Society.

Huang, X., & Reinisch, B. W. (2006). Real-time HF ray tracing through a tilted ionosphere. *Radio Science*, *41*(5). https://doi.org/10.1029/2005RS003378

Hysell, D. L. (2018). *Antennas and radar for environmental scientists and engineers*. Cambridge University Press.

International Association of Geomagnetism and Aeronomy. (2019). *International geomagnetic reference field*. https://www.ngdc.noaa.gov/IAGA/vmod/igrf.html

Australian Government. (2016). *Introduction to HF radio propaga-tion*. http://www.sws.bom.gov.au/Educational/5/2/2

Ishimaru, A. (2017). *Electromagnetic wave propagation, radiation, and scattering*. IEEE Press.

Jones, M. R., & Stephenson, J. J. (1975). *A versatile three-dimensional ray tracing computer program for radio waves in the ionosphere*. U.S. Dept. of Commerce.

Juras, A. S. (1985). *Handbook of geophysics and the space environment*. Air Force Geophysics Lab.

Lighthill, M. J. (1965). Group velocity. In M. Froissart (Ed.), *Hyper-bolic equations and waves* (pp. 96–123). Springer.

Misra, P., & Enge, P. (2011). *Global Positioning System*. Ganga-Jamuna Press.

National Geospatial Intelligence Agency. (n.d.). *World Geodetic System 1984 (WGS 84)*. https://www1.nga.mil/ProductsServices/GeodesyandGeophysics/WorldGeodeticSystem/Pages/default.aspx

Nickisch, L. J. (1988). Focusing in the stationary phase approx-imation. *Radio Science*, *23*(2), 171–182. https://doi.org/10.1029/RS023i002p00171

Nickisch, L., Fridman, S., Hausman, M., & Kraut, S. (2016). Assim-ilative modeling of ionospheric dynamics for nowcasting of HF propagation channels in the presence of TIDs. *Radio Science*, *51*(3), 184–193. https://doi.org/10.1002/2015RS005890

Nocedal, J., & Wright, S. J. (2006). *Numerical optimization*. Springer.

Psiaki, M. L. (2019). Ionosphere ray tracing of radio frequency sig-nals and solution sensitivities to model parameters. *Radio Science*, *54*(8), 738–757. https://doi.org/10.1029/2019RS006792

Psiaki, M. L., Ward, K. C., & DeMars, K. J. (2019). A bi-quintic latitude/longitude spline and lunar surface modeling for spacecraft navigation. *Journal of the Astronautical Sciences*, 1–47. https://doi.org/10.1007/s40295-019-00192-1

Psiaki, M. L., Ward, K. C., & DeMars, K. J. (2019). A bi-quintic lat-itude/longitude spline and lunar surface modeling for spacecraft navigation. *Journal of the Astronautical Sciences*, *67*, 1–47. https://doi.org/10.1007/s40295-019-00192-1

Psiaki, M. L., Bust, G. S., & Mitchell, C. N. (2015, September). Non-linear estimation to assimilate GPS TEC data into a regional iono-sphere model. *Proceedings of the 28th International Technical Meet-ing of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015)*, Tampa, FL, 3497–3510.

Reinisch, B. W., & Bilitza, D. (2004). Karl Rawer's life and the history of IRI. *Advances in Space Research*, *34*(9), 1845–1850. https://doi.org/10.1016/j.asr.2004.09.002

Seybold, J. S. (2005). *Introduction to RF propagation*. Wiley.

Stankov, S. M., Jakowski, N., Heise, S., Plamen, M., Kutiev, I., & War-nant, R. (2003). A new method for reconstruction of the vertical electron density distribution in the upper ionosphere and plasma-sphere. *Journal of Geophysical Research - Space Physics*, *108*(A5). https://doi.org/10.1029/2002JA009570

Yeh, K. C., & Liu, C. H. (1972). *Theory of ionospheric waves*. Academic Press.

Zaalov, N. Y., Moskaleva, E. V., & Burmakina, T. S. (2017). Application of the IRI model to the HF propagation model with optimization of the ionosphere parameters to day-to-day variation. *Advances in Space Reasearch*, *60*(10). https://doi.org/10.1016/j.asr.2017.08.018