Distributed ionosphere monitoring by collaborating mobile receivers

Okuary Osechas & Jason H. Rife
Tufts University

>> Accepted Article <<

CITATION:
O. Osechas and J. Rife (2014). Distributed ionosphere monitoring by collaborating mobile receivers, IEEE Transactions on Aerospace & Electronic Systems, 50(4):2860-2869. doi:10.1109/TAES.2014.120843. First published in Proc. ION Global Navigation Satellite Systems (ION GNSS 2012), Nashville, TN
DOI: 10.1109/TAES.2014.120843

FINANCIAL SUPPORT:
Federal Aviation Administration GBAS Program (Grant FAA-10-G-006)

Distributed Ionosphere Monitoring by Collaborating Mobile Receivers

Okuary Osechas, Jason Rife

Abstract

This paper presents a novel ionosphere anomaly monitoring method for Ground Based Augmentation Systems, based on a network of collaborative airborne dual-frequency receivers. The method allows aircraft to navigate safely using a single frequency or divergence-free position solution, methods more accurate than the ionosphere-free position solution. Each receiver measures ionospheric delays; measurements are shared among collaborators to estimate the local ionosphere gradient, which may be different for each ssatellite during a storm.

I. Introduction

Currently, the two most severe threats to Ground Based Augmentation Systems (GBAS) for navigation satellite systems are severe ionosphere storms and elevated levels of radio-frequency interference (RFI) [1]. Ionospheric disturbances are structures in the ionosphere that cause unusually large spatial gradients of ionosphere delay in the GPS signals. They are a threat to the integrity of GBAS, because if the corrections broadcast by GBAS contain wrong ionosphere information a solution will appear valid when it is, in fact, not [2]. Threats related to radio frequency interference (RFI) have been observed to be caused by unintentional disruption of GBAS reference stations through personal privacy devices (PPDs) [3]. RFI threats mostly represent a continuity threat; if reference receivers are disrupted GBAS becomes unavailable for a period of time, but it is unlikely that invalid information is broadcast.

Conventional wisdom suggests that the threat posed by ionospheric disruptions can be reduced by introducing dual-frequency measurements [4]. The ionospheric delay of a signal is inversely proportional to the square of its frequency, making it possible to estimate or remove the ionospheric delay using GNSS signals on two frequencies. Two types of measurement filtering

that would reduce or eliminate the integrity threat posed by severe ionosphere storm fronts are ionosphere-free and divergence-free processing [4]. Ionosphere-free processing can eliminate ionosphere delays, but at the expense of multiplying the size of nominal errors by a factor of nearly three [4], when compared to single-frequency processing. Divergence-free processing eliminates code-carrier divergence errors that accumulate over time, but it does not eliminate absolute ionosphere delay errors introduced by spatial separation between the mobile user and ground-station antennae.

It has been suggested that divergence-free processing will be advantageous for GBAS applications, to preserve high availability, but only if some form of supplemental monitoring is possible to mitigate the impact of severe storms [5]. In addition, it has been shown in [6] that the threefold increase in noise, associated with the ionosphere-free dual-frequency solution over the conventional single-frequency solution, is a factor which would degrade navigation availability severely enough to make GBAS unacceptable from an operational point of view.

Potential candidates for supplemental ionosphere-storm monitoring include a combination of geometry screening [7] and code-carrier divergence monitoring [8], the long baseline monitor [9], and the ionosphere gradient monitor (IGM) [10].

Cat-I operations were enabled, in large part, through the introduction of geometry-screening methods [7]. This approach inflates user protection levels by adjusting error models (e.g. sigma parameters) broadcast by the ground-station, such that users are rigorously protected against even worst-case ionosphere storm scenarios. The disadvantage of geometry screening is that the process of sigma-inflation reduces GBAS availability.

To mitigate availability penalties associated with geometry screening, many enhanced ionosphere monitors have been proposed for Cat-III GBAS (also known as GAST-D). The simplest of these monitors is a long-baseline monitor [9], which would monitor ionosphere gradients using a pair of ground-based receivers separated by ten kilometers or more. This approach is very difficult to implement at most airports, both because of the difficulty of siting and securing ground-based antennae off airport property and because of the difficulties of siting off-shore

monitor receivers for ocean-side airports.

A more compact but algorithmically complicated monitor, based on carrier-phase data, is called the Ionosphere Gradient Monitor (IGM) [10]. The IGM uses antennae spaced over short distances (hundreds of meters) to form carrier-phase differences that can be used to detect threatening ionosphere gradients. Because the IGM can be located on the airport surface, it is likely to be incorporated in GAST-D systems. Some of the shortcomings of this approach are (1) that IGM algorithms place severe restrictions on antenna siting in order to achieve good observability of ionosphere fronts and (2) that IGM installations are susceptible to elevated RFI, since antennas must be located in close proximity, typically within hundreds of meters.

This paper presents a novel method for ionosphere anomaly monitoring. It is based on a networked approach, where collaborating users share their estimates of ionosphere delays to collectively estimate the local ionosphere gradient. The magnitude of the gradient is used to detect anomalous conditions, when the ionospheric delay gradient becomes large enough to invalidate differential corrections broadcast by GBAS.

The proposed approach is shown to provide a useful tool in ensuring safe operation of GBAS, independent of satellite geometry. Accessing measurements from other receivers enables a user to compute a more reliable estimate of the state of the local ionosphere than a one-point measurement would. The proposed monitoring approach would enable high availability dual-frequency monitoring for severe ionosphere storms that might impact the integrity of users computing single-frequency or divergence-free position solutions.

Another important feature of the proposed monitor is that it would operate at a scale that is an intermediate between the local scale of GBAS and the regional scale of Space-Based Augmentation Systems (SBAS); as such it is able to capture ionosphere structures at a smaller scale than SBAS.

As compared to conventional monitoring approaches centered at a single GBAS facility, the proposed approach offers two key advantages: it extends monitoring capabilities over a much

wider area, and it permits continuing coverage even when a subset of monitor receivers are jammed by elevated RFI. Increasing the coverage area would be achieved by providing long-baseline monitoring across the entire region in which the GBAS VHF Data Broadcast can be received. By comparison, airport-based monitors typically only cover a limited region ($10\ km$ or fewer). Network-based monitoring has the potential to extend coverage out to the largest allowed coverage radius for GBAS (45 to $60\ km$). This effect is a compelling advantage of the proposed method, but it will not be discussed in detail in this paper.

Conventional ionosphere monitors may be disabled by severe RFI conditions that corrupt one or more GBAS reference stations; carrier-phase IGM is particular vulnerable to localized interference from PPDs, because IGM requires monitor antennas to be close together (within a few hundred meters) [11]. In a distributed monitoring approach, by contrast, GBAS ionosphere monitoring would remain active (and GBAS service might remain available), so long as differential corrections could be broadcast from at least one ground station antenna or from an alternative source [12]. This compelling advantage will be investigated using simulation later in the paper, after the proposed monitoring approach is described in detail in the coming sections.

II. System Architecture

The proposed system requires a set of aircraft equipped with dual-frequency receivers, as well as a GBAS ground station, also dual-frequency equipped. Using dual-frequency processing, each aircraft and the GBAS station are able to obtain measurements of ionospheric delay for each satellite. A data communications link is required for users to share the ionospheric delay measured for each satellite, as well as the position of the user receiver, itself (see Figure 1). Using compiled ionosphere delays for an individual satellite as seen by all users, the gradient affecting that satellite can be estimated. The estimated gradient on each satellite is monitored for extreme values that indicate the presence of an anomaly. If such an anomaly is detected, the monitor triggers an alert and declares the satellite unusable. Computation of the gradient estimates can happen either centrally, with all measurements concentrated in one location (e.g. the GBAS ground station), or it can be decentralized, with every user receiver computing its own estimates.

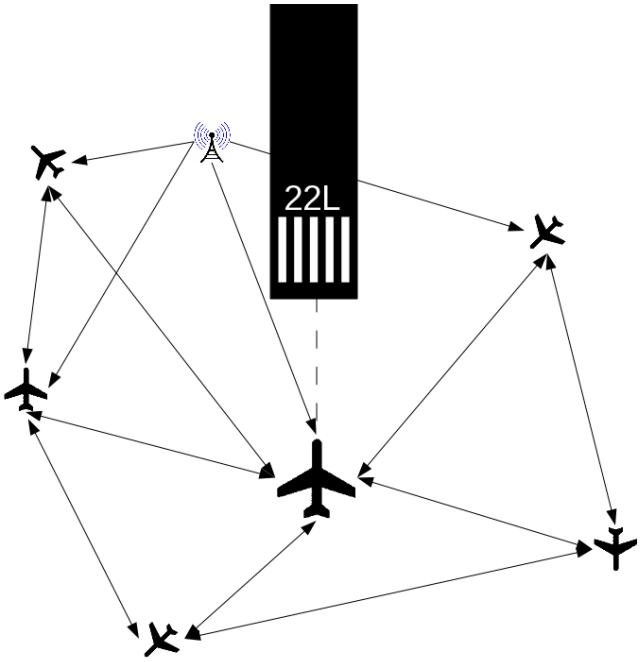

Fig. 1. The proposed system leverages a network of airborne receivers to probe the ionosphere at different locations and estimate local gradients.

III. Threat Model

The hazardous impact of severe ionosphere storms on GBAS has been a concern since moving ionosphere-delay fronts, featuring gradients as large as $400\ mm/km$, were identified by Datta-Barua et al. [2]. Typically, hazardous ionosphere storms are modeled as a traveling wedge; however, observations of anomalies with different structures have also been reported [13]. Figure 2 shows a set of different possible models for ionospheric anomalies; the wedge model is shown on the top, while the bottom two models correspond to so-called ionosphere "bubbles".

Using a network of dual-frequency receivers, ionospheric delay (I) can be measured by different receivers (i), at different locations ((x_i, y_i)) for different satellites (k), in order to compute a planar description of the local ionosphere using the thin-shell assumption (see [9]).

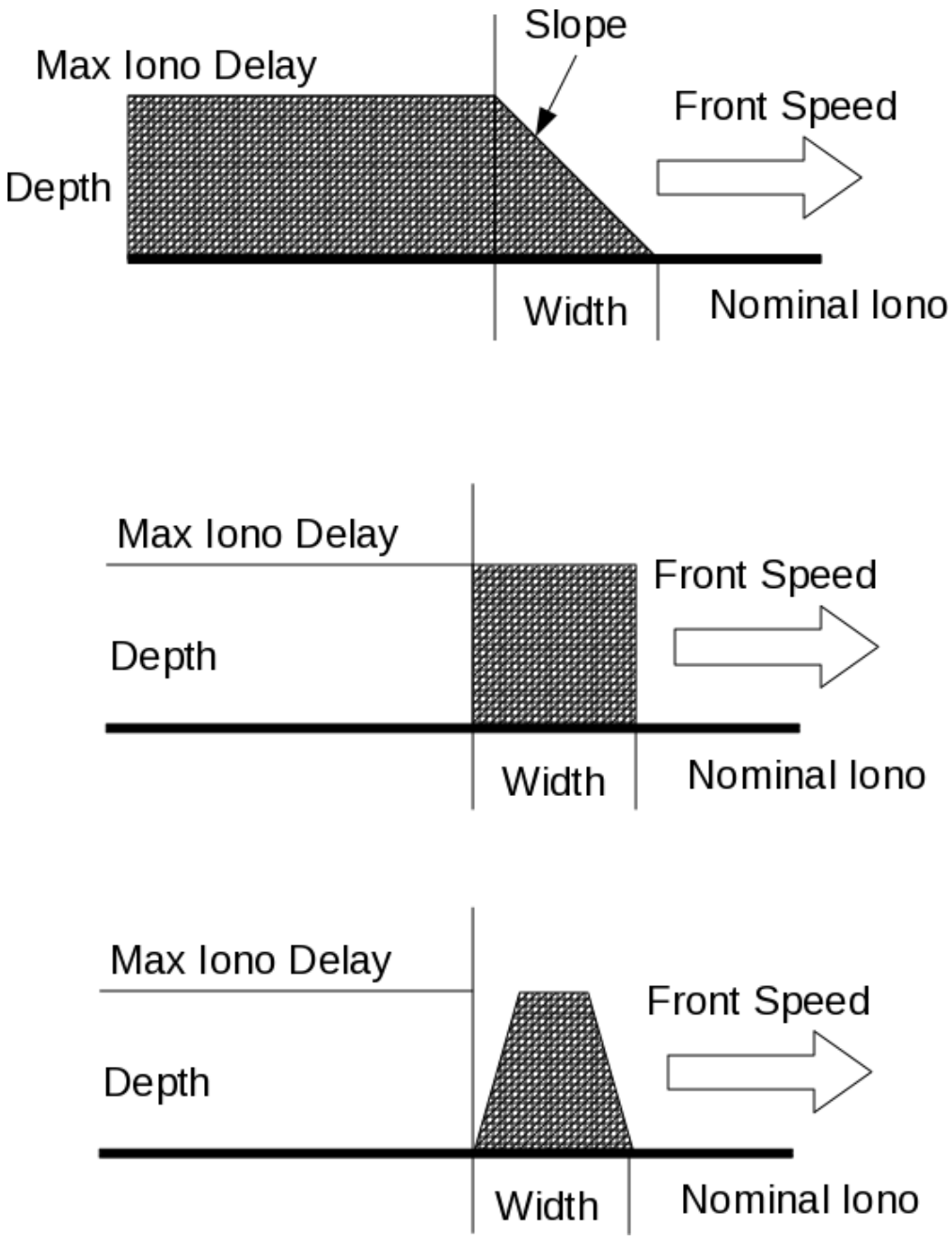

Fig. 2. Three different categories of ionospheric anomalies, adapted from [13].

The thin shell assumption is justified in that all receivers (airborne and ground based) are far below the altitude of the ionosphere.

We assume that, during severe ionosphere storms, ionosphere anomalies introduce disturbances that are highly correlated over regions of tens of kilometers across. As a corollary, we assume that anomalous storm activity will introduce salient gradients over local regions (defined as circles of 50 km in radius in this paper). Those local gradients are modeled as approximately planar. No other assumptions regarding the structure of an ionosphere storm are made in this paper. Because storm structure is modeled as being local in nature, a different ionosphere gradient should be computed for each particular satellite. Since the groups of pierce points corresponding to different satellites can be thousands of kilometers apart, no attempt is made to infer ionosphere structure between ionosphere pierce points (IPPs) for different satellites.

The ionosphere delay I for each satellite (j), as seen by user i, is estimated assuming a locally planar model for variations in ionosphere delay.

$$I_i^{(j)} = f(\mathbf{x}_i^{(j)}) = \mathbf{g}^{(j)^T}\mathbf{x_i^{(j)}} + c^{(j)}, \tag{1}$$

For this purpose we define the position vector $\mathbf{x}_i^{(j)} = \left[x_i^{(j)} y_i^{(j)}\right]^T$ and the vector of gradient parameters $\mathbf{g}^{(j)} = \left[a^{(j)} b^{(j)}\right]^T$.

Here $x_i^{(j)}$ and $y_i^{(j)}$ are the east and north coordinates of the pierce points of satellite (j) seen from receiver i, in a coordinate system centered at the GBAS reference. The variable $I_i^{(j)}$ is the measurement of ionospheric delay for the corresponding pierce point. Note that the inter-frequency bias for satellite k is assumed to be common for all users and is hence lumped into the parameter $c^{(j)}$.

IV. Proposed Method

The networked ionosphere monitor collects measurements of ionospheric delay from at least three receivers and computes a local ionosphere gradient estimate ($\hat{\mathbf{g}}^{(j)}$) for each satellite.

$$\hat{\mathbf{g}}^{(j)} = \begin{bmatrix} \hat{a}^{(j)} \\ \hat{b}^{(j)} \end{bmatrix} \tag{2}$$

If the magnitude of the estimated gradient is larger than a threshold, an alarm can be issued. This section describes the detailed mathematical formulation for the proposed ionosphere monitor.

A. Local Gradient Estimate for Each Satellite

As a simplifying assumption, the gradient (2) is computed under the assumption of a planar ionosphere, as described in equation (1). The measurements of ionosphere delay for all users can be compiled in the following matrix form:

$$\begin{bmatrix} \vdots \\ I^{(j)}(x_i, y_i) \\ \vdots \end{bmatrix} = \left[\begin{array}{cc|c} \vdots & \vdots & \vdots \\ x_i^{(j)} & y_i^{(j)} & 1 \\ \vdots & \vdots & \vdots \end{array}\right] \left[\begin{array}{c} a^{(j)} \\ b^{(j)} \\ \hline c^{(j)} \end{array}\right] \tag{3}$$

For simplicity of notation, the vector of measurements $I^{(j)}(x_i, y_i)$ is denoted as $\mathbf{y}^{(j)}$. The matrix of $(x_i^{(j)}, y_i^{(j)})$ positions is named $\mathbf{A}_g^{(j)}$, and the vector of all ones is named $\mathbf{A}_c^{(j)}$. The estimated parameters $\hat{a}^{(j)}$ and $\hat{b}^{(j)}$ are grouped into the gradient vector $\hat{\mathbf{g}}^{(j)}$. Equation (3) can now be re-written as:

$$\mathbf{y}^{(j)} = \left[\begin{array}{c|c} \mathbf{A}_g^{(j)} & \mathbf{A}_c^{(j)} \end{array}\right] \left[\begin{array}{c} \mathbf{g}^{(j)} \\ \hline c^{(j)} \end{array}\right] = A \left[\begin{array}{c} \mathbf{g}^{(j)} \\ \hline c^{(j)} \end{array}\right]. \tag{4}$$

The local ionosphere gradient (or local planar fit for ionospheric delay) can be estimated via $\mathbf{A}^+$, the weighted pseudo-inverse of $\mathbf{A}$.

$$\left[\begin{array}{c} \hat{\mathbf{g}}^{(j)} \\ \hline \hat{c} \end{array}\right] = \mathbf{A}^+ \left(\mathbf{y}^{(j)} + \boldsymbol{\nu}^{(j)}\right) \tag{5}$$

with $\mathbf{A}^+$ being equal to $\left(\mathbf{A}^T\ \mathbf{W}\ \mathbf{A}\right)^{-1}\ \mathbf{A}^T\ \mathbf{W}$, with $\mathbf{W} = \mathbf{R}^{-1}$ being the inverse of the sensor noise covariance (defined below), and with $\boldsymbol{\nu}$ being the measurement noise vector.

Assuming the measurement noise distribution is zero-mean and Gaussian, the distribution of the gradient estimate $\hat{\mathbf{g}}$ is also Gaussian. Given a general form of a multi-variate Gaussian is

$$\Phi\left(\hat{\mathbf{g}}; \mathbf{g}, \mathbf{Q}\right) = \frac{1}{|2\pi\mathbf{Q}|^{\frac{1}{2}}} exp\left(-\frac{1}{2}\left(\mathbf{z} - \boldsymbol{\mu}\right)^T \mathbf{Q}^{-1}\left(\mathbf{z} - \boldsymbol{\mu}\right)\right), \tag{6}$$

with the distribution of the estimated gradient:

$$p(\hat{\mathbf{g}}) = \Phi(\hat{\mathbf{g}}; \mathbf{g}, \mathbf{Q}_{est}). \tag{7}$$

The mean of the $\hat{\mathbf{g}}$ distribution is the actual gradient vector $\mathbf{g}$. The covariance of the estimate is $\mathbf{Q}_{est}$ defined as:

$$\mathbf{Q}_{est} = E\left[\mathbf{g}'^{(j)} \left(\mathbf{g}'^{(j)}\right)^T\right], \tag{8}$$

where $\mathbf{g}'^{(j)}$ is the estimation error, defined as:

$$\mathbf{g}'^{(j)} = \mathbf{P}\mathbf{A}^{+}\boldsymbol{\nu}^{(j)}. \tag{9}$$

Since $c^{(j)}$ is not used in the monitoring statistic, the projection matrix $\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ is needed to separate the terms of interest from the estimation equation (5). Thus,

$$\mathbf{Q}_{est} = \mathbf{P}\mathbf{A}^{+}\mathbf{R}\mathbf{A}^{+^T}\mathbf{P}^T. \tag{10}$$

B. Monitoring Statistic

The threat to GBAS is proportional to the gradient magnitude. Thus, in order to tie monitor performance to gradient magnitude, it is useful to make the monitoring statistic m equal to the estimate of this quantity.

$$m = \|\hat{\mathbf{g}}^{(j)}\|^2 = (\hat{\mathbf{g}}^{(j)})^T\ \hat{\mathbf{g}}^{(j)}. \tag{11}$$

The distribution of the monitoring statistic is given by cummulative of the density function, defined in (6) to describe the distribution of the ionosphere gradient:

$$p(m) = \int_0^m \Phi\left(\hat{\mathbf{g}}; \mathbf{g}, \mathbf{Q}_{est}\right) \tag{12}$$

Since the estimation errors are not necessarily identically distributed in both coordinate directions, the distribution of m is not χ^2 but *generalized* χ^2. Designing thresholds for generalized χ^2 distributions and computing MDEs requires the implementation of specialized numerical algorithms, as discussed in [14], [15]. This generalized χ^2 distribution can be used to evaluate the distribution of the monitoring statistic in the presence of a fault with a particular gradient $\mathbf{g}$.

Whereas (12) is useful for characterizing fault scenarios, a slight reformulation is necessary in determining the monitor threshold, to account for nominal variations in $\mathbf{g}$.

The matrix $\mathbf{Q}_{env}$ can be assumed to be made up of identical diagonal entries, which is equivalent to assuming ionospheric disturbances have no particular direction of preference:

$$\mathbf{Q}_{env} = \begin{bmatrix} \sigma_I^2 & 0 \\ 0 & \sigma_I^2 \end{bmatrix}. \tag{13}$$

A model for the long-term distribution of the monitor statistic, which accounts for the variability of the ionosphere over periods of many hours, is also generalized χ^2, with the following distribution:

$$p_{long}(m) = \int_0^m \Phi\left(\hat{\mathbf{g}}; \mathbf{0}, \mathbf{Q}_{est} + \mathbf{Q}_{env}\right) d\hat{\mathbf{g}} \tag{14}$$

C. Monitor Threshold

Under nominal operations, i.e. when no ionospheric anomaly is present, the monitor should issue as few warnings as necessary to maximize availability. The monitor threshold T can be computed to ensure that the integral of the long-term monitor distribution above T does not exceed the continuity allocation for the monitor ($p_{FA,req}$). The probability to be computed is that of triggering an alarm, given that no fault is present:

$$p_{long}(m > T|\text{fault-free}) \leq p_{FA,req}. \tag{15}$$

D. Faulted operations

The monitoring statistic should minimize the probability of not detecting a hazardous ionospheric anomaly. To meet GBAS integrity requirements, the probability that no alarm is triggered when a fault is present should not exceed the maximum allowed probability of a missed detection ($p_{MD,req}$). Note that this is an instantaneous requirement, rather than a long-term average, and so the distribution in (12) is used to evaluate the missed-detection risk:

$$p(m \leq T|\text{fault present}) \leq p_{MD,req}. \tag{16}$$

In order to assess monitor sensitivity, we assess the left side of the above equation, defined as the missed-detection probability p_{MD}, as a function of gradient size $\|\mathbf{g}\|$. We will call this a

"p_{MD} vs. G" analysis, to evoke the connection to the "p_{MD} vs. E" approach, which has become a standard tool in the analysis of GBAS-related systems [16].

V. Simulation Scenario

A computer simulation was set up to demonstrate the method and study its benefits and potential shortcomings, as compared to existing or proposed methods. In this simulation, airborne receivers are placed uniformly at random in 2-D space, within a certain radius (currently set at $50\ km$) over San Francisco Bay. In concept we imagine that a GBAS ground facility is located at an airport at the center of this circle. Noisy measurements of the ionosphere are aggregated over a number of receivers to yield an estimate of the current ionosphere gradient, thus providing a capability to detect anomalies injected into the simulated ionosphere. A snapshot of one such set of 19 airborne receivers can be seen in Figure 3.

The ionosphere measurements (the entries of $\mathbf{y}^{(j)}$, in equation (4)) are assumed to have additive noise with standard deviation $\sigma_i = 1\ m$. The noise on each receiver is assumed to be independent from other receivers, which leads to a diagonal sensor covariance matrix $\mathbf{R}$. The nominal ionosphere is also assumed to have a diagonal covariance matrix $\mathbf{Q}_{env}$ as defined by (13); where $\sigma_I = 4\ x\ 10^{-6} = 4\ mm/km$, which is in accordance with [2].

Different user-receiver geometries directly impact the measurement covariance matrix $\mathbf{Q}_{est}$, which in turn sets the detection thresholds and the probability of mis-detecting a fault as a function of fault magnitude (16). For instance, the p_{MD} vs. G curve for a set of 19 receivers randomly distributed over a circular area of $50\ km$ in radius is shown in figure Figure 4. In the specific case of Figure 4, the detection threshold was set to a $113\ mm/km$. This particular geometry achieves a p_{MD} of 10^{-9} when the anomalous gradient reaches $131\ mm/km$.

VI. Dependence of Sensitivity on Receiver Density

An important feature of the proposed method is that increasing the number of participants improves the sensitivity. This happens as increasing the number of receivers both enhances

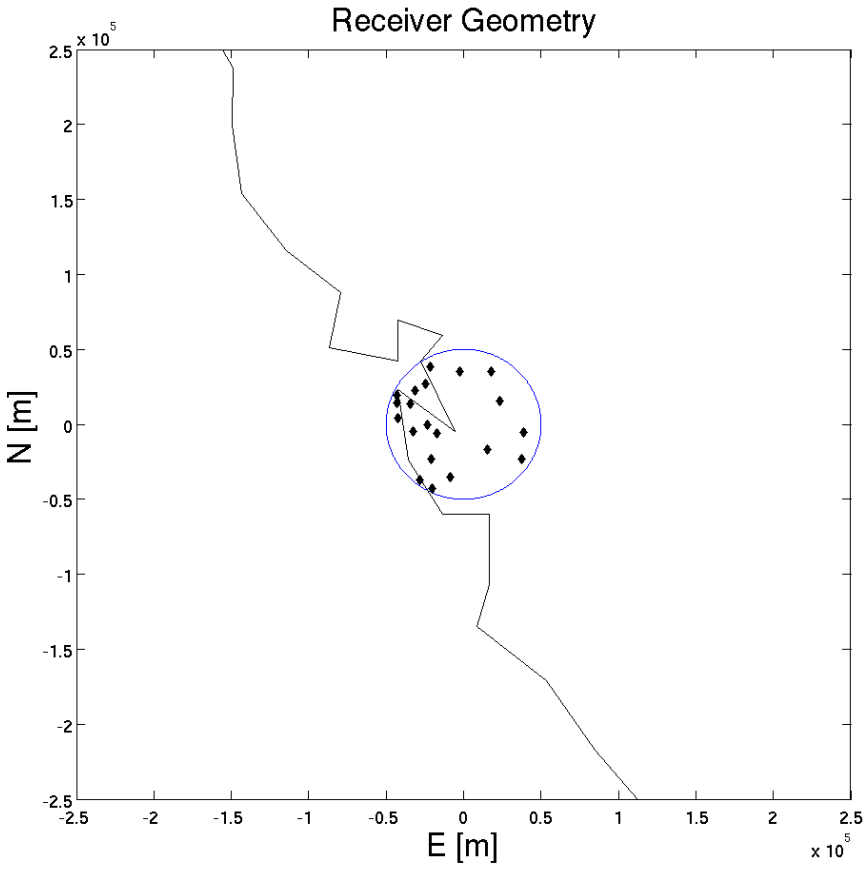

Fig. 3. Network of 19 simulated airborne receivers over Oakland, CA. Note that a radius of 50 km around Oakland airport covers two further international airports: San Francisco International Airport and San Jose International Airport.

geometric diversity and provides enhanced averaging of noisy signals, thus shrinking the ellipse associated with the matrix $\mathbf{Q}_{est}$. In the limit of a large number of user receivers, this, in turn, leaves the ionospheric variability ($\mathbf{Q}_{env}$) as the dominant term setting the threshold T according to (15).

In order to study the sensitivity as a function of receiver density, different numbers of simulated receivers were distributed uniformly within the $50\ km$ communications radius. For each configuration of receivers the covariance matrix $\mathbf{Q}_{est}$ was calculated, and from it the detection threshold. Although a full integrity verification would consider a P_{MD} vs. G curve of the form

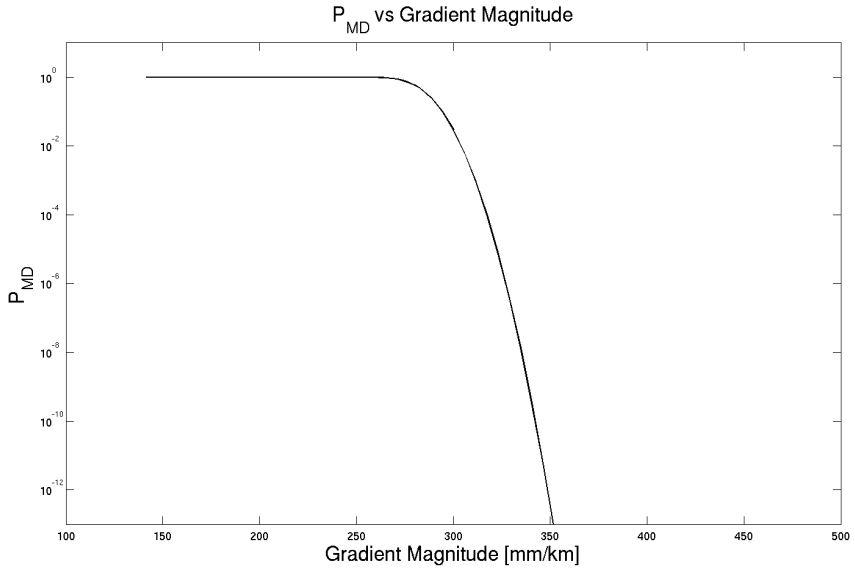

Fig. 4. p_{MD} vs E behavior of the receiver distribution seen in Figure 3

of Figure 4, it is useful in this study to characterize sensitivity using a single number, the value on the P_{MD} vs. G curve that corresponds to a P_{MD} value of 10^{-9}. We label this point the minimum detectable gradient (MDG). The MDG is the smallest value of $\|\mathbf{g}\|$ that can reliably be deteced. Using simulation, MDG was computed for 72 randomly generated geometries for each of 20 different numbers of collaborators between 8 and 800. The results of these simulations are plotted in Figure 5. The graph indicates the mean MDG for each set size, and the standard deviation from the mean. The variability observed in the plot can be attributed to the fact that the different geometries are generated at random. For the geometries simulated, this variability was small relative to the MDG, and so the mean values are a good representation of overall performance.

The graph in Figure 5 shows an approximately linear dependence between the logarithm of the MDG and the logarithm of the number of collaborating receivers. It is evident from looking at the plot that increasing the number of users in the network will increase the sensitivity without requiring longer baselines. When a large number of receivers collaborate (beyond the right side of Figure 5), the missed-detection probability asymptotes to a constant value determined only

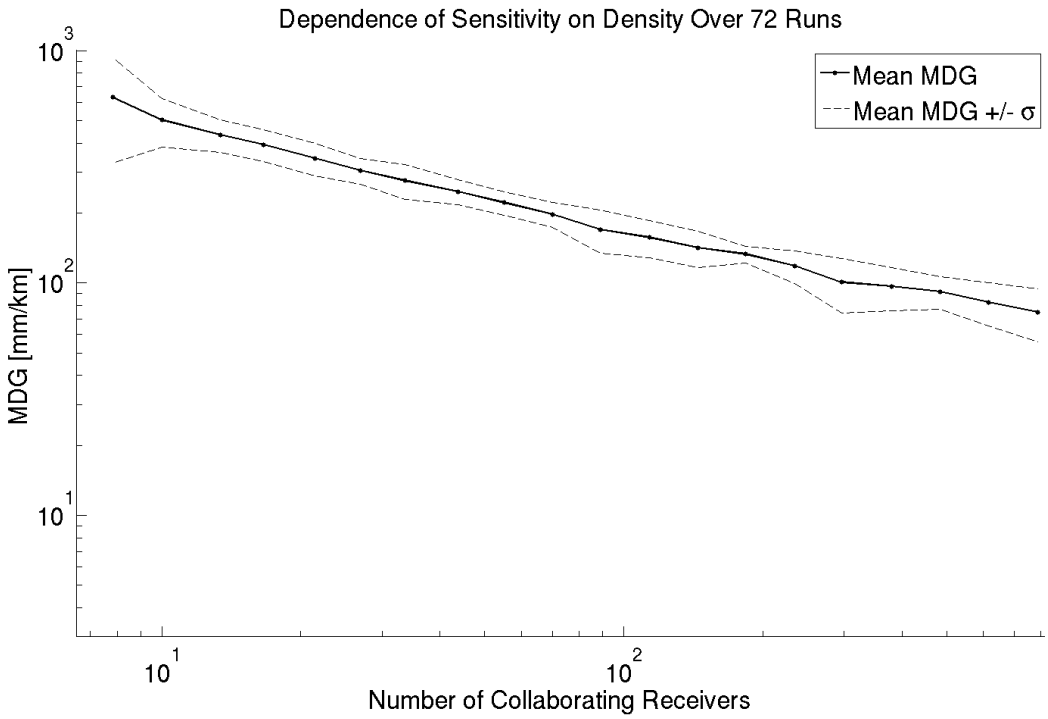

Fig. 5. The sensitivity of the method depends on the density of collaborating receivers.

by the nominal ionosphere variations with covariance ($\mathbf{Q}_{env}$).

VII. RFI Robustness of the Method

This section discusses the benefits of the proposed collaborative ionosphere monitor with regards to its robustness to radio-frequency interference (RFI).

In order to illustrate this point, a simulation experiment was applied to study how the sensitivity of the proposed method would be impacted if users in the vicinity of the GBAS facility were excluded from the solution. In this scenario, it is assumed that differential corrections are available from another source, perhaps from a second GBAS reference station located in the same urban metroplex.

In the RFI simulation 200 simulated users were placed, uniformly distributed within the $50\ km$

communications radius. One receiver at the center of the circle represents the ground facility. A hypothetical RFI source (or PPD) was also situated at the origin; this RFI source was assumed to deny service to any participating users within the interference radius. In our simulations we allow for an extended jamming radius, extending from as little as $0\ km$ to as much as $49\ km$. The denied users were simply not considered in calculating the covariance matrix ($\mathbf{Q}_{est}$). The resulting degradation in p_{MD} vs G behavior for different interference radii is depicted in Figure 6.

Figure 6 shows that the algorithm provides robust detection of ionosphere anomalies, with little change to sensitivity out to a jamming of $20\ km$. In fact, sensitivity does not degrade significantly until the jamming radius exceeds $45\ km$. It is important to keep in mind that users are maximally $50\ km$ away from the origin, which means that at $49\ km$ nearly 96% of the coverage area has been disabled. This kind of situation far more severe than the unintentional RFI scenarios described in [3].

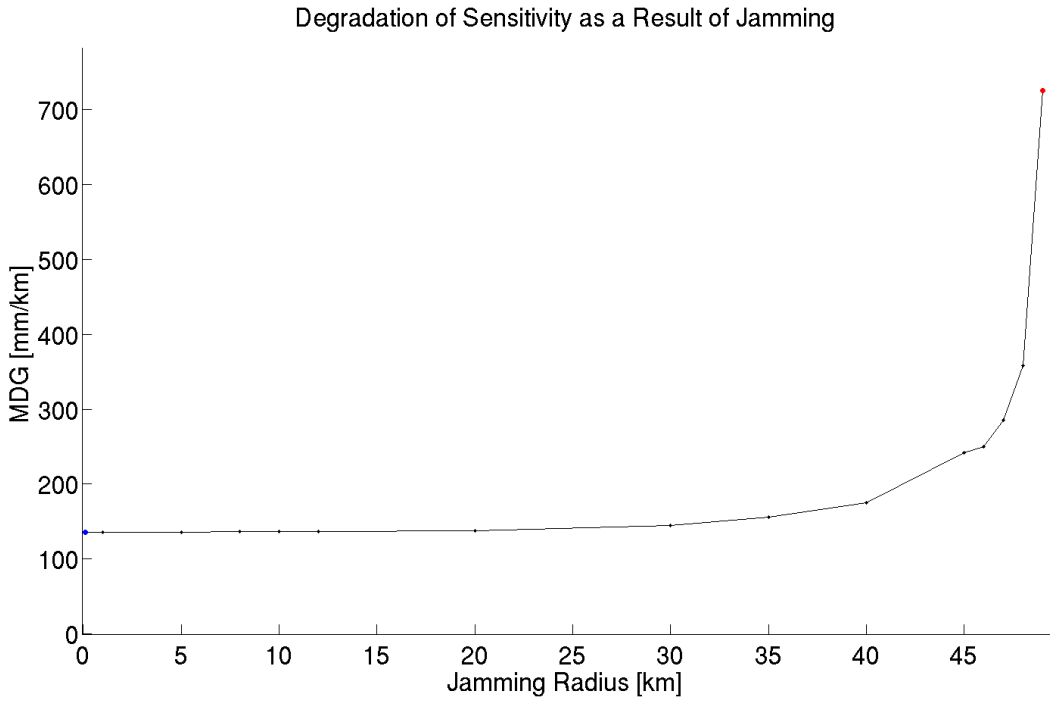

Fig. 6. Influence of interference radius on the MDG. The blue dot (to the left end) represents the nominal case of a $0\ km$ interference radius, while the red dot (to the right end) represents a radius of $49\ km$.

VIII. Operation Under Restricted Bandwidth

As the RFI simulations show, the proposed algorithm performs robustly even when only a subset of available measurements are used for monitoring. This characteristic of the algorithm suggests that it might also perform robustly in a bandwidth-limited communication scenario, in which the available communication channel cannot broadcast all user measurements for inclusion in the monitor statistic. This section will consider in more detail this limited-bandwidth scenario, by proposing and analyzing a receiver-selection algorithm designed to maximize monitor sensitivity given practical communications constraints.

The analysis of sensitivity as a function of the number of users, presented in section VI, does not take into consideration the restrictions potentially imposed by bandwidth limitations in the communications network. Under a scenario of limited bandwidth, it is no longer feasible to assume that data are available from all N users, and it becomes necessary to define a criterion for selecting a subset of k receivers when computing the monitoring statistic m.

Selecting an optimal subset of size k from the set (Ω_N) of all users N can be thought of as an optimization problem. Within Ω_N there are $\begin{pmatrix} N \\ k \end{pmatrix}$ different subsets (Ω_k^i), each with k users ($k \leq N$). The problem becomes one of selecting the subset (Ω_k^*) of receivers that will yield the smallest MDG. This can be formulated as an optimization problem:

$$\Omega_k^* = \text{argmin}\left(MDG(\Omega_k^i)\right). \tag{17}$$

Given that there are a finite number of different Ω_k^i, a global optimum is guaranteed. Finding this global optimum can, however, become computationally intensive as N and k grow. This combinatorial explosion may hinder the scheme from working in a real-time situation. In order to mitigate the combinatorial explosion, this section discusses two relaxations to the optimization problem that make it tractable to solve in real-time. The resulting tradeoff between reduction of computing time and monitor sensitivity are subsequently discussed.

A. Relaxations to the Optimization Problem

This section proposes two relaxations to the optimization problem: 1) a greedy search algorithm to replace the exhaustive combinatorial search, and 2) an approximate cost function that is faster to compute than the MDG. The decrease in the sensitivity of the scheme due to both relaxations is shown to be negligibly small compared to the decrease in computation time.

1) Greedy Search Algorithm: The greedy search is based on the assumption that a desirable, yet sub-optimal, set of $k+1$ receivers (Ω'_{k+1}) contains the previously found desirable set of k receivers (Ω'_k). This relaxation eases the computational requirement involved in finding the optimal set of k receivers, as it reduces the number of different subsets of receivers that need to be evaluated for an adequate geometry to be found. In that sense, for some given initial number of participants, k_0, the desired set of receivers is defined recursively with

$$\Omega'_{k+1} = \text{argmin}\left(\text{MDG}(\Omega^i_{k+1})\right), \text{ such that } \Omega'_k \subset \Omega'_{k+1} \tag{18}$$

The initialization of the scheme requires a selection of an initial set of k_0 receivers. For a small k_0, it is computationally inexpensive to pick the initial set Ω'_{k_0} using an exhaustive search of $\begin{pmatrix} N \\ k_0 \end{pmatrix}$ combinations of k_0 receivers using (17). Thus, $\Omega'_{k_0} = \Omega^*_{k_0}$.

Once an initial set of receivers has been determined, the greedy search algorithm adds receivers to the estimate one at a time. Every time a new receiver is to be added, all candidates are evaluated by adding them to a geometry matrix and computing either the MDG, or a proxy cost function, described in the next section. In either case, the heuristic reduces the number of cost function evaluations to $N-k$, rather than $\begin{pmatrix} N \\ k \end{pmatrix}$.

2) Proxy Cost Function: Computation of the MDG, as defined in equation (16) for $P_{MD} = 10^{-9}$, is the most computationally expensive element of the algorithm. One way to speed up the calculation of the optimal subset of receivers is to find a metric that will yield similar results, but is faster to compute than MDG.

Reducing the MDG is equivalent to increasing the geometric diversity of the set of receivers;

intuitively this corresponds to enlarging the baselines between sensors. A different, surrogate metric for geometric diversity is the second moment (or statistical covariance) of the user receiver positions. This second moment may be visualized as an ellipse in which larger ellipse dimensions in any direction correspond to larger baselines in that direction. Since sensitivity is enhanced by larger baselines, it is desirable to maximize the principal axes of the second-moment ellipse. An equivalent problem is to minimize the negative of the smallest principal axis. Define $\Sigma_{A_g^{(j)}}$ as the central second-moment (or covariance) of the set of user positions contained in the geometry matrix $\mathbf{A}_g^{(j)}$. The cost function then becomes:

$$J = -min(eig(\Sigma_{A_g^{(j)}})). \tag{19}$$

The eigenvalues of $\Sigma_{A_g^{(j)}}$ indicate the principal axes of the error distribution; the estimate is most sensitive along the direction of the eigenvector associated with the largest eigenvalue. Thus, the cost function J penalizes small eigenvalues, as this encourages the selection of Ω_k^* as the subset that is most sensitive along its least sensitive direction.

The combination of both relaxations to the optimization problem finally yields an expression for the recursive definition of the subset Ω'_k, given an initial subset Ω^*_{K0}. The recursion becomes:

$$\Omega'_{k+1} = \text{argmin}\left(\text{J}(A_g^{(j)}(\Omega'_k))\right), \text{ such that } \Omega'_k \subset \Omega'_{k+1}. \tag{20}$$

B. Impact of the Relaxations

The objective of the two relaxations, previously discussed, is to make the optimization problem tractable. The performance of the relaxed optimization method was analyzed in a simulation in which the two methods of selecting the desired subset of receivers were compared for different values of k. The subsets generated from the combination of greedy search and proxy cost function, Ω'_k as defined in equation (20), were compared the true optimum Ω_k^*, generated from the exhausitve search for the minimization problem of equation (17).

Despite being a suboptimal approach, the proposed method performs remarkably well, as shown in Figure 7. The plot shows the difference in detection performance between the baseline sets for each k and the greedy search heuristic. Figure 7 was generated for one geometric

configuration of receivers.

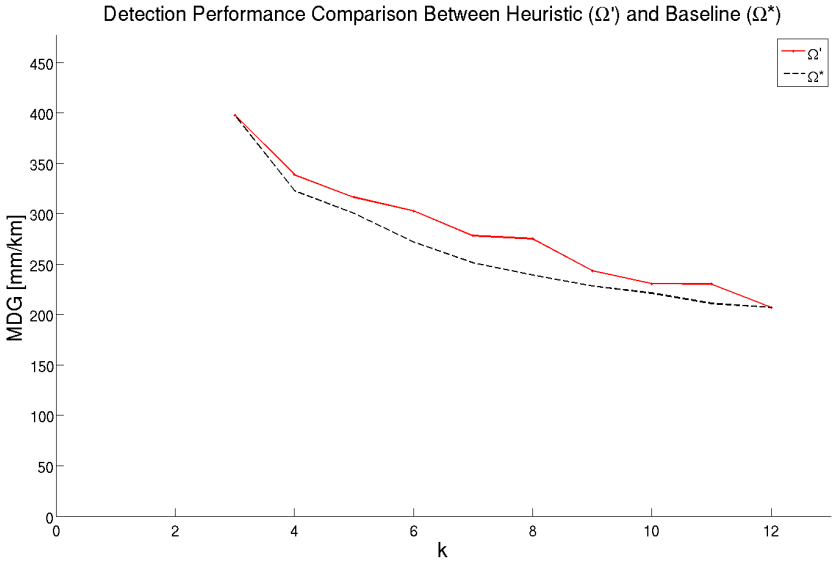

Fig. 7. Difference in MDG between the greedy algorithm and the global optimum. Computed for $k_0 = 3$, $N = 10$

Note that the difference in MDG is zero for $k = k_0$ as well as for $k = N$. For $k = N$ this is expected, as there is only one possibility for choosing N in N; for $k = k_0$, the fact that the sensitivities are identical is a confirmation that the proxy cost function J, defined in equation (19), is appropriate, as in these simulations it selects the same optimal seed set as the full MDG computation does.

The benefit of the greedy algorithm is that the number of evaluations of the equation (16) can be drastically reduced for $3 < k < N - 3$. Recall that the method is not defined for $k < 3$ and that the number of combinations in $\begin{pmatrix} N \\ k \end{pmatrix}$ peaks around $k = N/2$ but drops off as k approaches N. The number of evaluations M_{eval} required for the thorough search of the most sensitive geometry is $\begin{pmatrix} N \\ k \end{pmatrix}$, while the greedy search algorithm requires M_{eval} of $\begin{pmatrix} N \\ k_0 \end{pmatrix}$ for

$k = k_0$ and

$$M_{eval} = \begin{pmatrix} N \\ k_0 \end{pmatrix} + \sum_{i=k_0+1}^{N} (N - i) \text{ for } k > k_0. \quad (21)$$

A comparison of the computing time required for each scheme is shown in Figure 8. Considering that for some $k = 5$ the baseline case requires up to $5\ s$ to be computed, but it takes less than $40\ ms$ to compute the heuristic, it is evident from Figure 8 that there is a gain in processing time for the heuristic approach that can be attributed to the two relaxations discussed above, except for the case when $k = N$. Note that the case $k = N$ is trivial as there is only one possible Ω_N and no search is required.

An important aspect of the heuristic search algorithm is that the benefits increase as N increases. The complexity of the optimal search is combinatorial, while the complexity of the heuristic is linear with N. The results of Figure 8 were computed for a representative case, where the effect is noticeable even for only 12 users. In an equivalent graph over 200 users, the gain in computation time is significantly larger. As an illustration of the reduction in complexity due to the greedy search algorithm, consider that there are $\begin{pmatrix} 200 \\ 18 \end{pmatrix} = 1.8913\ x\ 10^{25}$ different combinations of 18 receivers in a set of 200; By comparison, the number of evaluations required by the greedy algorithm is 19 orders of magnitude lower: $1.316\ x\ 10^6$.

The proposed heuristics do not significantly degrade monitor sensitivity. The decrease in sensitivity that is traded for a gain in computation time is addressed in Figure 9. The data underlying this plot were generated from 10 repetitions of the experiment discussed in Figure 7; for each repetition, the difference in MDG between the baseline and the heuristic is normalized by the baseline MDG. The plot shows that as k increases, the degradation of the MDG increases slightly before decreasing. Even at the worst k values, the MDG only degraded by 17% (mean) or $\approx 29\%$ (maximal value over 10 trials). Thus, the scheme works better for larger k, but even for a bad k the mean decrease in MDG is below 20%

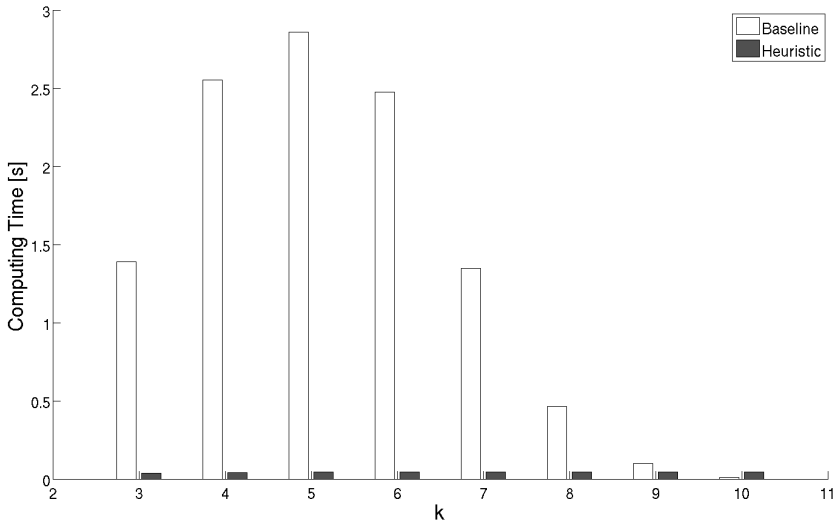

Fig. 8. A comparison of the computing time required for calculating the optimal (Ω_k^*) and near optimal (Ω_k') subsets of participants for different set sizes (k). For the plotted example, the heuristic search requires an initial computing time of about 36 ms for $k = 3$, with increasing computing times as k grows; this increase, however is small compared to the first step and becomes progressively smaller. The increase from $k = 3$ to $k = 4$ is of 6 ms, but only 0.3 ms from $k = 9$ to $k = 10$).

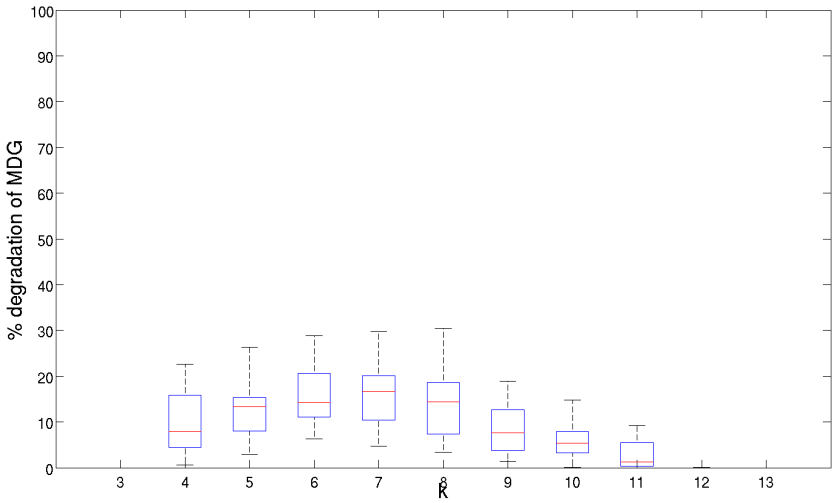

Fig. 9. Percentual degradation of MDG computed for the sets selected by the greedy algorithm, as compared to the set with best sensitivity for each k. Results computed over 10 different trials.

IX. Discussion

This section discusses some of the benefits of distributed ionosphere monitoring over other methods; such advantages include the absence of siting requirements and robustness to RFI. The limitations of the method, in the form of communication requirements, are also discussed.

A. Benefits of the Method

A first benefit is that the networked ionosphere gradient monitor, unlike IGM, has no explicit siting constraint related to integer ambiguity. By comparison the standard IGM method imposes strict siting requirements that limit reference receivers to be positioned within a few hundred meters of one another [11].

A second additional benefit of networked ionosphere gradient monitoring is the possibility of enabling a relaxation of D_{max}. The current value of D_{max} (approximately $10\ km$) means that services can only be offered to aircraft within that distance of the ground facility. The current restriction is derived from performance limitations of local ground-station-based ionosphere gradient monitors [17], rather than limitations on the range of the data broadcast. With a mechanism in place for detecting widespread ionospheric anomalies, a relaxation of D_{max} to the limits of reliable communication becomes possible. This means that, if RFI can be otherwise mitigated, the cost-benefit advantages of using one GBAS reference station to cover multiple airports in an urban metroplex might be achieved.

A potential non-aviation benefit of networked ionosphere monitoring is to support collaborative automotive applications. With vehicle-to-vehicle communications becoming more and more feasible, co-operative navigation schemes become more attractive. The proposed collaborative ionosphere gradient monitor would complement other collaborative methods for verifying GNSS integrity, such as CERIM [18]. Deploying an automobile-based variant of the proposed scheme would benefit from the fact that automobiles are able to operate in a configuration that is significantly more dense than for aircraft.

B. Limitations of the Method

The single most prominent weakness of the method is the dependence on aircraft-to-aircraft data communication. While such systems have been discussed in the scope of the NextGen program, under the System Wide Information Management (SWIM), there is no indication of this capability coming into operation soon. Alternatively, a modified ADS-B message might be the quickest means for enabling the proposed networked ionosphere monitoring concept.

X. Summary

This paper introduces a novel ionosphere monitor based on a collaborative network of airborne dual-frequency receivers. Ionosphere delay measurements are aggregated across multiple receivers to compute an estimate of the local ionosphere gradient for each satellite. These estimates are then used to form a monitoring statistic to detect anomalously large ionosphere gradients.

Simulations were used to assess the sensitivity of the proposed algorithm under several scenarios. In the first scenario the influence of the number of receivers on the sensitivity was analyzed. A second scenario looks at the robustness of the method when RFI removes a subset of collaborators, and the third scenario shows how the scheme can be implemented under limited bandwidth. The simulation scenarios show consistently that, with a sufficient number of collaborators, the scheme is able to detect ionospheric gradients of $200\ mm/km$ with a probability of missed detection below 10^{-9}.

XI. Acknowledgments

The authors gratefully acknowledge the Federal Aviation Administration GBAS Program (Grant FAA-10-G-006) for supporting this research. The opinions discussed here are those of the authors and do not necessarily represent those of the FAA or other affiliated agencies.

References

[1] Global Navigation Satellite System Program Office, "Gnss evolutionary architecture study. phase i - panel report," Tech. Rep., Federal Aviation Administration, 2008.

[2] Seebany Datta-Barua, Jiyun Lee, Sam Pullen, Ming Luo, Alexandru Ene, Di Qiu, Godwin Zhang, and Per Enge, "Ionospheric Threat Parameterization for Local Area Global-Positioning-System-Based Aircraft Landing Systems," *Journal of Aircraft*, vol. 47, pp. 1141–1151, 2010.

[3] Sam Pullen and Grace Gao, "GNSS jamming in the name of privacy," *Inside GNSS*, March/April 2012.

[4] Patrick Y. Hwang, Gary A. McGraw, and John R. Bader, "Enhanced differential gps carrier-smoothed code processing using dual-frequency measurements," *NAVIGATION : Journal of The Institute of Navigation*, vol. 46, no. 2, 1999.

[5] Hiroyuki Konno, Sam Pullen, Jason Rife, and Per Enge, "Evaluation of Two Types of Dual-Frequency Differential GPS Techniques under Anomalous Ionosphere Conditions," in *ION ITM*, 2006.

[6] Hiroyuki Konno, "Dual-Frequency Smoothing for CAT III LAAS: Performance Assessment Considering Ionosphere Anomalies," in *ION GNSS*, 2007.

[7] Jiyun Lee, Jiwon Seo, Young Shin Park, Sam Pullen, and Per Enge, "Ionospheric threat mitigation by geometry screening in ground-based augmentation systems," *JOURNAL OF AIRCRAFT*, vol. 48, no. 4, 2011.

[8] D.V. Simili and B. Pervan, "Code-carrier divergence monitoring for the gps local area augmentation system," in *Position, Location, And Navigation Symposium, 2006 IEEE/ION*, 25-27, 2006, pp. 483 – 493.

[9] Ming Luo, Sam Pullen, Todd Walter, and Per Enge, "Ionosphere spatial gradient threat for laas: Mitigation and tolerable threat space," in *ION National Technical Meeting*, 2004.

[10] Samer Khanafseh, Sam Pullen, and John Warburton, "Carrier Phase Ionospheric Gradient Ground Monitor for GBAS with Experimental Validation," *NAVIGATION: Journal of The Institute of Navigation*, vol. 59, no. 1, 2012.

[11] Jing Jing, S. Khanafseh, Fang-Cheng Chan, S. Langel, and B. Pervan, "Detecting ionospheric gradients for gbas using a null space monitor," in *Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION*, april 2012, pp. 1125 –1133.

[12] J. Rife, "Collaborative Vision-Integrated Pseudorange Error Removal: Team-Estimated Differential GNSS Corrections with no Stationary Reference Receiver," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 13, no. 1, pp. 15 –24, march 2012.

[13] Matt Harris and Tim Murphy, "GAST D Ionosphere Anomaly Simulation Including Airplane Speed Change Effect on Monitors," Tech. Rep., NSP July 09 CSG/Flimsy 2, July 2009.

[14] Jason Rife, "Overbounding Chi-square Probability Distributions," in *Proc. ION GNSS 2012*, 2012.

[15] Mathieu Joerger, Fang-Cheng Chan, and Boris Pervan, "Addressing the Limitations of Existing RAIM Methods by Fully Exploiting Measurement Redundancy," in *Proc. ION GNSS 2012*, 2012.

[16] Young C. Lee and Curtis A. Shively, "Feasibility of GPS III integrated with an inertial system to provide CAT IIIb services," in *Proc. of ION GNSS 17th International Technical Meeting of the Satellite Division*, 2007, pp. 267 – 284.

[17] Ming Luo, Sam Pullen, Alexandru Ene, Di Qiu, Todd Walter, and Per Enge, "Ionosphere threat to laas: Updated model, user impact, and mitigations," in *ION GNSS*, 2004.

[18] J. Rife, "Collaboration-enhanced receiver integrity monitoring with common residual estimation," in *Proc. IEEE/ION Position, Location, and Navigation Symposium (PLANS 2012)*, 2012.