Autonomous Measurement Outlier Detection and Exclusion for Ground Vehicle Navigation with Cellular Signals and IMU

Mahdi Maaref and Zaher M. Kassas, Senior Member, IEEE

Abstract—An autonomous measurement outlier detection and exclusion framework for ground vehicle navigation using cellular signals of opportunity and an inertial measurement unit (IMU) is developed. In this framework, the ground vehicle aids its onboard IMU with cellular pseudoranges in a tightly-coupled fashion in the absence of global navigation satellite system (GNSS) signals. First, cellular pseudorange measurements are characterized from an extensive wardriving campaign collected with a ground vehicle in different environments: open sky, urban, and deep urban. Then, a framework is developed, which accounts for sources of measurement outliers: (i) line-of-sight blockage and (ii) short multipath delays. These outliers induce unmodeled biases in the cellular pseudorange measurements, which compromise the integrity of the navigation solution. Simulation results are presented for a ground vehicle navigating with proposed receiver autonomous integrity monitoring (RAIM)-based framework in the presence of measurement outliers. It is demonstrated that the proposed framework could autonomously detect and exclude measurement outliers, reducing the position root mean squared error (RMSE) by 88.56% and the maximum position error by 79.09%. Experimental results are presented evaluating the efficacy of the proposed framework on a ground vehicle navigating in a deep urban environment over trajectory of 1500 m in the absence of GNSS signals. The experimental results demonstrate the proposed framework successfully detecting and excluding outlier measurements, reducing the position RMSE by 41.5% and the maximum position error by 43.1%.

Index Terms—Outlier detection, Cellular, RAIM, GNSS, IMU, navigation

I. INTRODUCTION

GLOBAL navigation satellite system (GNSS) signals are insufficient for reliable and accurate ground vehicle navigation in deep urban environments due to the inherent weakness of their space-based signals. Hong Kong’s dense urban environment, for instance, has less than 50% reliable GPS coverage [1]. What is more, GNSS signals are susceptible to malicious cyber attacks in the form of spoofing and jamming, which is particularly threatening as ground vehicles move towards semi- and full-autonomy [2]–[4].

Alternative sensing modalities to compensate for GNSS limitations and vulnerabilities have been the subject of extensive research recently. Such sensors include inertial measurement units (IMUs) [5], [6], lidars [7], [8], cameras [9], radars [10], [11], among others. Signals of opportunity represent another particularly fruitful class of sensing modalities [12]–[15]. Signals of opportunity are radio frequency signals that are not intended for navigation but can be exploited for navigation purposes, such as AM/FM radio signals [16], [17], digital television [18], [19], cellular [20], [21], and low Earth orbit (LEO) satellites [22]–[24]. Among these signals, cellular signals are particularly attractive due to their (1) abundance in urban environments, (2) favorable geometric configuration by construction of the cellular infrastructure, (3) high received power due to their proximity, (4) transmission at various frequencies, and (5) high transmission bandwidth. These characteristics render them a reliable and accurate alternative navigation source, which is inherently more difficult to jam and spoof. Recent research results have demonstrated meter-level accurate navigation with cellular signals of opportunity with ground vehicles [25]–[28] and centimeter-level accurate navigation with aerial vehicles [29], [30].

In safety-critical applications, such as ground transportation and aviation, it is imperative to assess the integrity of the navigation solution. Integrity monitoring of GNSS signals and measurement outlier detection and exclusion have been extensively studied in the literature [31]–[34]. Among the different methods that have been developed to monitor GNSS signal integrity, receiver autonomous integrity monitoring (RAIM) inherently possesses desirable characteristics for ground-based receivers in urban canyons, due to RAIM’s design flexibility and adaptability to urban environments [35], [36]. In contrast to Satellite Based Augmentation System (SBAS) and Ground Based Augmentation System (GBAS) integrity methods, RAIM alleviates the need for costly, bulky, and computationally intensive infrastructure. RAIM detects GNSS pseudorange measurement outliers by only exploiting the redundancy of GNSS signals to check the measurements’ consistency. RAIM can also be coupled with other sensing modalities to enhance the systems’ integrity [37].

Despite the promise of signals of opportunity as a reliable and accurate sensing modality, their integrity has not been fully studied. This paper develops an autonomous measurement outlier detection and exclusion method for cellular long-term evolution (LTE) signals of opportunity. An initial work on integrity monitoring for cellular LTE signals-based navigation was conducted in [38], where the integrity of the LTE-based navigation solution and its corresponding protection levels (PLs) were studied. This paper extends [38] and makes two
main contributions. First, outliers in received cellular LTE pseudorange measurements are characterized, namely biases due to line-of-sight (LOS) signal blockage and biases due to short multipath delays. Second, an autonomous measurement outlier detection and exclusion approach is developed and validated numerically and experimentally. This paper considers the following practical scenario. A ground vehicle is equipped with a GNSS receiver, an IMU, and a receiver capable of producing pseudoranges to multiple cellular LTE towers in the environment. The vehicle has initial access to GNSS signals, which are fused with the IMU measurements to estimate the vehicle’s states (orientation, position, velocity, IMU biases, and GNSS receiver clock error). GNSS signals become unusable as the vehicle navigates in the environment (e.g., as it enters a deep urban canyon or in the presence of a malicious jamming attack on the GNSS frequency band). In the absence of IMU aiding with GNSS signals, the IMU errors grow unboundedly. However, cellular signals can be used as an IMU aiding source in the absence of GNSS signals to bound the IMU errors [39], [40]. Given the low elevation angle of cellular towers, cellular signals inevitably experience LOS blockage and suffer from multipath, which introduce unmodeled biases in the estimated pseudorange, rendering these pseudorange measurements outliers, which would contaminate the vehicle’s estimated position, compromising the navigation system’s integrity. This paper develops an autonomous outlier detection and exclusion framework for cellular LTE signals experiencing LOS blockage and multipath. The proposed framework assumes the position of the cellular towers to be known a priori. In addition, the proposed framework assumes the presence of a stationary agent in the vehicle’s environment, referred to as the base, which has knowledge of its own state at all time. The base’s purpose is to estimate the dynamic stochastic clock bias states of the cellular transmitters and to share these estimates with the navigating vehicle. This navigating vehicle uses the cellular pseudoranges as an aiding source to correct the IMU’s drift in the absence of GNSS signals. If the vehicle is equipped with other navigation sensors, adding pseudoranges from cellular towers via the radio mapping (e.g., [41], [42]) or satellite images. It is also possible to generate a priori (e.g., from a local or a cloud-hosted database). This database could be generated a priori via several approaches, such as radio mapping (e.g., [41], [42]) or satellite images. It is also assumed that the vehicle has an initial period of access to GNSS signals. During this period, the vehicle estimates its state and the cellular transmitters’ clock bias states. After this period, it is assumed that GNSS signals become unusable, and the vehicle begins to navigate exclusively with cellular signals and its onboard IMU. The remainder of this section describes the vehicle-mounted receiver clock error dynamics, vehicle’s kinematic model, and the EKF-based navigation framework.

II. NAVIGATION FRAMEWORK

This section describes the ground vehicle navigation framework. The environment is assumed to comprise $N_s$ terrestrial cellular transmitters, denoted $\{S_n\}_{n=1}^{N_s}$. It is assumed that the vehicle knows the location of the cellular transmitters (e.g., from a local or a cloud-hosted database). This database could be generated a priori via several approaches, such as radio mapping (e.g., [41], [42]) or satellite images. It is also assumed that the vehicle has an initial period of access to GNSS signals. During this period, the vehicle estimates its state and the cellular transmitters’ clock bias states. After this period, it is assumed that GNSS signals become unusable, and the vehicle begins to navigate exclusively with cellular signals and its onboard IMU. The remainder of this section describes the vehicle-mounted receiver clock error dynamics, vehicle’s kinematic model, and the EKF-based navigation framework.

A. Receiver Clock Error Dynamics Model and EKF Time Update

The receiver clock error state is given by

$$x_{clk,r} \triangleq [c\delta t_r, c\dot{\delta} t_r]^T,$$

where $c$ is the speed of light, $\delta t_r$ is the receiver’s clock bias, and $\dot{\delta} t_r$ is the receiver’s clock drift. The vehicle-mounted receiver clock error state is assumed to evolve according to a double integrator driven by process noise $\bar{w}_{clk,r} \triangleq [\bar{\dot{\delta} t}_{r,t}, \bar{\ddot{\delta} t}_{r,t}]^T$, whose elements are modeled as zero-mean, mutually independent noise processes and the power spectral density of $\bar{u}_{clk,r}$ is given by $Q_{clk,r} = \text{diag} \{S_{\dot{\delta} t}_{r,t}, S_{\ddot{\delta} t}_{r,t}\}$ [43]. The discrete-time equivalent of the clock error dynamics can be expressed as

$$x_{clk,r}(k+1) = \Phi_{clk} x_{clk,r}(k) + w_{clk,r}(k), \quad \Phi_{clk} \triangleq \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \quad (1)$$

where $k$ is the measurement time-step, $T$ is the sampling time, and $w_{clk,r}$ is the process noise, which is modeled as a zero-mean white random sequence with covariance

$$Q_{clk,r} = \sigma^2 \begin{bmatrix} S_{\dot{\delta} t}_{r,t} T + S_{\dot{\dot{\delta} t}_{r,t}} T^2 & S_{\dot{\delta} t}_{r,t} \dot{T} \\ \dot{T} S_{\dot{\delta} t}_{r,t} & \dot{T}^2 S_{\dot{\dot{\delta} t}_{r,t}} \end{bmatrix}. \quad (2)$$

Note that the power spectra $S_{\dot{\delta} t}_{r,t}$ and $S_{\dot{\dot{\delta} t}_{r,t}}$ can be related to the power spectral density of the fractional frequency deviation $y(t)$ of an oscillator from nominal frequency, which has the form $S_y(f) = \sum_{\alpha=-2}^{\alpha=2} h_\alpha f^\alpha$. A common approximation
of $S_{\delta t,r}$ and $S_{\delta t,r}$, is one that only considers the white frequency coefficient $h_0$ and the frequency random walk coefficient $h_{-2}$, specifically, $S_{\delta t,r} \approx \frac{h_0}{\tau}$ and $S_{\delta t,r} \approx 2\pi^2 h_{-2}$ [44], [45].

Given the model in (1), the EKF time update can be readily calculated to yield the predicted state estimate $\hat{x}_{clk,r}(k+1|j)$, for $j \leq k$.

B. Vehicle Kinematics Model and EKF Time Update

The vehicle is assumed to be equipped with:

- an IMU and
- a receiver capable of producing pseudorange measurements to cellular transmitters (e.g., [19], [21], [27], [28], [46, 47]).

If the vehicle is equipped with other navigation sensors (e.g., lidar, camera, etc.) the proposed framework could seamlessly integrate the outputs of these sensors to improve the vehicle’s navigation solution. The vehicle’s state vector $x_v$ is defined as

$$x_v \triangleq \begin{bmatrix} \dot{G}q^T, G\dot{r}_r, \dot{b}_g^T, \dot{b}_a^T \end{bmatrix}^T,$$

where $l_Gq$ is the unit quaternion representing the vehicle’s orientation (i.e., rotation from the global frame $G$ to the IMU’s inertial frame $I$); $G\dot{r}_r \triangleq [x_r, y_r, z_r]^T$ and $G\dot{r}_r$ are the 3-D position and velocity of the vehicle, respectively, expressed in $G$; and $\dot{b}_g$ and $\dot{b}_a$ are the gyroscope and accelerometer biases, respectively. The IMU noisy measurements of angular rotation $l_G\omega_r$ and linear acceleration $G\dot{a}_r$, corrupted by the gyroscope and accelerometer biases, respectively. The IMU samples the angular rate $\omega_{\text{meas}}$ and specific forces $a_{\text{meas}}$ every $T$ seconds and can be modeled as

$$\omega_{\text{meas}}(k) = l_G\omega_r(k) + b_g(k) + n_g(k),$$

$$a_{\text{meas}}(k) = R[l_G\dot{q}]\begin{bmatrix} G\dot{r}_r(k) \end{bmatrix} + b_a(k) + n_a(k),$$

where $I_k$ is the inertial frame at time-step $k$, $R[\dot{q}]$ is the equivalent rotation matrix of $\dot{q}$, $G\dot{q}$ is the acceleration due to gravity in the global frame, and $R\dot{q}$ and $n_{\text{g}}$ and $n_a$ are measurement noise vectors, which are modeled as zero-mean white noise sequences with covariances $\sigma_{\omega}^2 I_{3\times3}$ and $\sigma_{\text{a}}^2 I_{3\times3}$, respectively. The orientation of the IMU evolves according to

$$l_G^{k+1}q = I_k \otimes l_k \dot{q},$$

where $\otimes$ is the quaternion multiplication operator and $l_{k+1} \dot{q}$ represents the relative rotation of the inertial frame from time-step $k$ to $k+1$, which is the solution to the differential equation

$$l_{k} \dot{q} = \frac{1}{2} \Omega [l_{\omega}(\tau)] l_{k} \dot{q}, \quad \tau \in [t_k, t_{k+1}],$$

where $t_k \triangleq kT$ and for any vector $\omega \triangleq [\omega_1, \omega_2, \omega_3]^T \in \mathbb{R}^3$, the matrix $\Omega[\omega]$ is defined as

$$\Omega[\omega] \triangleq \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}.$$  \hspace{1cm} (5)

The velocity evolves according to

$$G\dot{r}_r(k+1) = G\dot{r}_r(k) + \int_{t_k}^{t_{k+1}} G\alpha_r(\tau) \, d\tau.$$ \hspace{1cm} (6)

The position evolves according to

$$G\hat{r}_r(k+1) = G\hat{r}_r(k) + \int_{t_k}^{t_{k+1}} G\dot{r}_r(\tau) \, d\tau.$$ \hspace{1cm} (7)

The biases $b_g$ and $b_a$ will be modeled as random walk processes, i.e.,

$$\dot{b}_g = w_g, \quad \dot{b}_a = w_a,$$ \hspace{1cm} (8)

where $w_g$ and $w_a$ are modeled as zero-mean white random processes with covariances $\sigma_{\omega}^2 I_{3\times3}$ and $\sigma_{\text{a}}^2 I_{3\times3}$, respectively. Additional details of the orientation, position, and velocity models are discussed in [48], [49]. Given the model (8), the EKF time update can be readily calculated to yield $b_g(k+1|j)$ and $b_a(k+1|j)$, for $j \leq k$.

The orientation state estimate propagation equation is given by [49]

$$I_k^{k+1}q = I_k \hat{q} \otimes I_k \hat{q},$$ \hspace{1cm} (9)

where $I_{k+1} \hat{q}$ represents the estimated relative rotation of the vehicle from time-step $k$ to $k+1$. The expression of $I_k \hat{q}$ can be found in [49].

The vehicle’s velocity state estimate time update equation is obtained using trapezoidal integration and is given by [49]

$$G\hat{r}_r(k+1|j) = G\hat{r}_r(k|j) + \frac{T}{2} [\hat{s}(k|j) + \hat{s}(k+1|j)] + T G\hat{g}(k),$$ \hspace{1cm} (10)

where $\hat{s}(k|j) \triangleq R[q]k \hat{a}(k|j)$, $\hat{a}(k|j) \triangleq a_{\text{meas}}(k) - b_a(k|j)$, $R[q]k \hat{a}(k|j)$ is the equivalent rotation of the vehicle from time-step $k$ to $k+1$. The expression of $I_k \hat{q}$ can be found in [49].

$$G\hat{r}_r(k+1|j) = G\hat{r}_r(k|j) + \frac{T}{2} [G\hat{r}_r(k+1|j) + G\hat{r}_r(k|j)].$$ \hspace{1cm} (11)

C. EKF Prediction Error Covariance Time Update

Subsections II-A and II-B described how to obtain the EKF state estimate time update for the vehicle-mounted receiver clock errors and vehicle’s quaternion, position, velocity, gyroscope bias, and accelerometer bias. This subsection discusses calculation of the prediction error covariance time update. The EKF estimates the state vector $x$ consisting of the vehicle’s and the receiver’s clock error states, i.e., $x \triangleq [x_r^T, x_{clk,r}^T]^T$.

$$\hat{x}(k|j) \triangleq \left[I_{k+1} q^T, G\hat{r}_r(k|j), G\hat{r}_r(k|j), b_g^T(k|j), b_a^T(k|j), x_{clk,r}(k|j) \right]^T,$$

the state estimate produced by the EKF at time-step $k$ obtained using all measurements (IMU and cellular pseudorange) from time-step 1 to $j \leq k$.

Subsequently, the EKF prediction error covariance time update is given by

$$P(k+1|j) = FP(k|j)F^T + Q,$$

$$F \triangleq \text{diag} [\Phi_r, \Phi_{clk,r}], \quad Q \triangleq \text{diag} [Q_r, Q_{clk,r}],$$
\[ \Phi_r = \begin{bmatrix} I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & \Phi_{gb} & \Phi_{gb} \\ \Phi_{rq} & I_{3 \times 3} & I_{3 \times 3}^T & \Phi_{gb} & \Phi_{gb} \\ \Phi_{rq} & 0_{3 \times 3} & I_{3 \times 3} & \Phi_{gb} & \Phi_{gb} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & I_{1 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & I_{1 \times 3} \end{bmatrix} , \]

\[ \Phi_{rq} = -\frac{T}{2} [\hat{s}(k,j) + \hat{s}(k+1,j)] \times , \quad \Phi_{rq} = \frac{T}{2} \Phi_{rq} , \]

\[ \Phi_{gb} = -\frac{T}{2} \sum \Phi_{gb} , \quad \Phi_{gb} = -\Phi_{gb} , \]

It is worth noting that a non-rotating global frame is assumed in the IMU measurement models. The detailed derivations of \( P(k|j) \), \( \Phi_r \), and \( Q_r \) are described in [48], [50].

D. EKF Measurement Model

After discretization and mild approximations, the pseudorange error process noise power spectra, i.e., \( \sigma_{\delta t_n,s}^2 \) and \( \sigma_{\delta t_n,r}^2 \), respectively. Denote \( c\delta t_{n,s}(k|k) \) as the KF’s \( n \)-th cellular transmitter clock bias estimate, \( c\delta t_{n,s}(k|k) \equiv c\delta t_{n,s}(k) - c\delta t_{n,s}(k) \) as the estimate error, and \( \sigma_{\delta t_n,s}^2(k|k) \) as its estimation error variance.

2) Pseudorange Measurement Model Used in the EKF:

The estimate of the cellular tower clock bias states produced by the base are sent to the navigating vehicle. Therefore, the pseudorange made by the vehicle-mounted receiver on the \( n \)-th cellular transmitter can be re-expressed as

\[ z_{n}(k) = \| G_{r} r(k) - r_{s,n} \|_2 + c \cdot [\delta t_{n}(k) - \delta t_{n,s}(k)] + v_{n}(k), \]

where \( v_{n}(k) \) is the measurement noise, which is modeled as a zero-mean white Gaussian sequence with variance \( \sigma_{ps,n}^2 \). The clock bias states of the cellular transmitters \( \{\delta t_{n}\}_{n=1}^{N} \) are assumed to be known to the navigating vehicle through a base receiver as discussed in Section II. 1) Cellular Transmitter Clock Bias Estimation:

3) Cellular Transmitter Clock Bias Estimation:

It is assumed that a base receiver is deployed in the same cellular environment of the navigating vehicle. In contrast to the ground vehicle, the base has complete knowledge of its position \( r_B \) and its clock bias \( c\delta t_{B}(k) \) for all \( k \), e.g., from GNSS measurements. The base is making the following pseudorange measurements to the \( n \)-th cellular transmitter in the vehicle’s environment

\[ z_{B,n}(k) = \| r_B - r_{s,n} \|_2 + c \cdot [\delta t_{B}(k) - \delta t_{n,s}(k)] + v_{B,n}(k), \]

where \( v_{B,n} \) is the measurement noise, which is modeled as a zero-mean white Gaussian random sequence with variance \( \sigma_{ps,n}^2 \). Since \( r_{s,n}, r_B, \) and \( c\delta t_{B}(k) \) are known, define the new measurement

\[ z'_{B,n}(k) \equiv \| r_B - r_{s,n} \|_2 + c\delta t_{B}(k) - z_{B,n}(k) \]

where \( \delta t_{s,n} \) is the clock drift of the \( n \)-th cellular transmitter. The dynamics of \( x_{clk,s,n} \) are the same as the ones of \( x_{clk,r} \) defined in (1), except that \( S_{\delta t_s,t} \) and \( S_{\delta t_r,t} \) are now replaced with the \( n \)-th cellular transmitter-specific process noise power spectra, i.e., \( S_{\delta t_s,t} \) and \( S_{\delta t_r,t} \), respectively.
rigorously define what constitutes an outlier in cellular pseudorange measurements. The next subsection studies cellular pseudorange measurements collected over hours of driving in different environments, leading to a statistical model of their accuracy. Next, outliers in cellular pseudoranges are analyzed in terms of biases due to LOS blockage and short multipath delays. Finally an outlier detection test is formulated.

A. Cellular Pseudorange Measurements Characterization

To analyze the accuracy of cellular LTE pseudorange measurements, a ground vehicle was driven for several hours in different environments: semi-urban, urban, and deep urban. For comparison purposes, pseudorange measurements from an antenna placed on the roof of the Engineering Gateway building at the University of California, Irvine were collected over 250 seconds to resemble an open sky environment. Received LTE signals from different LTE towers corresponding to different U.S. cellular providers transmitting at different frequencies and different cell-specific reference signal (CRS) bandwidths were collected, from which pseudorange measurements were obtained. The pseudoranges were obtained using the Multichannel Adaptive TRansceiver Information eXtractor (MATRIX) software-defined receiver (SDR) discussed in [27, 28]. Table I summarizes the characteristics of recorded cellular signals.

<table>
<thead>
<tr>
<th>Environment</th>
<th>$N_s$</th>
<th>Freq. [MHz]</th>
<th>Bandwidth [MHz]</th>
<th>Date (DD/MM/YYYY)</th>
<th>Traversed path [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open sky</td>
<td>2</td>
<td>2125 and 1955</td>
<td>20</td>
<td>13/5/2019</td>
<td>Stationary receiver</td>
</tr>
<tr>
<td>Semi-urban and urban</td>
<td>2</td>
<td>2145 and 1955</td>
<td>20</td>
<td>15/11/2016</td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2145 and 1955</td>
<td>20</td>
<td>20/1/2017</td>
<td>2000</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2145 and 739</td>
<td>20 and 10</td>
<td>22/6/2018</td>
<td>1600</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1955 and 739</td>
<td>20 and 10</td>
<td>27/6/2016</td>
<td>580</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>739</td>
<td>10</td>
<td>5/11/2017</td>
<td>825</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1955 and 739</td>
<td>20 and 10</td>
<td>22/8/2018</td>
<td>1800</td>
</tr>
<tr>
<td>Deep urban</td>
<td>2</td>
<td>2145 and 1955</td>
<td>20</td>
<td>12/10/2018</td>
<td>345</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2145 and 1955</td>
<td>20</td>
<td>24/8/2018</td>
<td>500</td>
</tr>
</tbody>
</table>

Next, the collected pseudorange measurements were compared against the true ranges to characterize the pseudorange measurement errors. Recall from (12) that the pseudorange is composed of three terms: (i) true range between the receiver and the transmitter (namely, $\|G_r(k) - r_{s_n}\|_2$), (ii) an error term due to the difference between the clock bias of the receiver and the transmitter (namely, $c\Delta\delta t_{r,s_n}(k) \triangleq c \cdot (\delta t_r(k) - \delta t_{s_n}(k))$), and (iii) an error term due to measurement noise (namely, $v_{s_n}(k)$).

To evaluate the statistics of the measurement noise term, $v_{s_n}(k)$, the true range between the receiver and the transmitter can be readily removed from knowledge of the cellular transmitters’ location and the receiver’s ground truth position (which is obtained from a GNSS receiver), i.e.,

$$z_{s_n}’(k) \triangleq z_{s_n}(k) - \|G_r(k) - r_{s_n}\|_2 = c\Delta\delta t_{r,s_n}(k) + v_{s_n}(k).$$

Next, the clock error term $c\Delta\delta t_{r,s_n}(k)$ must be removed. In GNSS, the satellites’ clock biases are transmitted in the navigation message and can be removed by the receiver from the pseudorange measurements. To reduce the receiver’s clock bias and drift issues, an atomic clock can be used [53]. Removing the effect of clock biases to evaluate the empirical measurement error of LTE pseudoranges is a more challenging problem. On one hand, the LTE transmitters’ clock biases, which are typically relatively stable [56] (e.g., using a GPS-disciplined oscillator (GPSDO)), are unknown to the receiver. On the other hand, the measured pseudoranges are produced by a receiver which does not possess a clock of atomic stability.

To remove $c\Delta\delta t_{r,s_n}(k)$ from $z_{s_n}’(k)$, it is assumed that this term evolves according to a first-order polynomial model with a constant initial clock bias $c\delta t_{r,s_n,0}$ and drift $c\delta t_{r,s_n,0}$ [57, 58], i.e.,

$$c\Delta\delta t_{r,s_n}(k) = c\delta t_{r,s_n,0} kT + c\delta t_{r,s_n,0} + \eta_{r,s_n}(k), \quad (16)$$

where $\eta_{r,s_n}(k)$ is a random sequence which models the mismatch between the true time evolution of $c\Delta\delta t_{r,s_n}(k)$ and the first-order polynomial approximation (16).

The constants $c\delta t_{r,s_n,0}$ and $c\delta t_{r,s_n,0}$ can be estimated via least-squares by post-processing the recorded data $z_{s_n}’(k)$. It is important to note that this is a conservative approach to characterize the pseudorange measurement error, as this approach calculates the term $\tilde{v}_{s_n}(k) \triangleq v_{s_n}(k) + \eta_{r,s_n}(k)$, which is larger than the true measurement error $v_{s_n}(k)$.

To analyze the accuracy of the pseudorange measurement error, the statistics of the term $\tilde{v}_{s_n}(k)$ will be characterized. Fig. 1(a) shows $z_{s_n}’(k)$ for one of the towers from the open sky recorded signals. The constants $c\delta t_{r,s_n,0}$ and $c\delta t_{r,s_n,0}$ are estimated from $z_{s_n}’(k)$ via least-squares and then subtracted from $z_{s_n}’(k)$ to get

$$z_{s_n}”(k) \triangleq z_{s_n}’(k) - \left[ c\delta t_{r,s_n,0} kT + c\delta t_{r,s_n,0} + \tilde{v}_{s_n}(k) \right].$$

Fig. 1(b) shows $z_{s_n}”(k)$ and Fig. 1(c)-(d) show the autocorrelation function (acf) and power spectral density (psd) of $z_{s_n}”(k)$, from which it can be seen that the term $\tilde{v}_{s_n}(k)$ is nearly white.

Therefore, if the receiver’s or transmitter’s clock type is unknown (i.e., the error component $\eta_{r,s_n}(k)$ cannot be modeled), the conservative term $z_{s_n}”(k)$ can be studied to characterize the pseudorange measurement error statistics.

In the case that both the receiver’s and transmitter’s clock types are known, the mismatch between the true time evolution
of \( c ∆ t_{r,s_n}(k) \) and its first-order polynomial approximation, \( η_{r,s_n}(k) \), can be characterized as follows. Using (1), the discrete-time equivalent dynamics model of the receiver and the transmitter clock error difference can be expressed as

\[
Δ x_{clk,r,s_n}(k + 1) = \Phi_{clk} Δ x_{clk,r,s_n}(k) + w_{clk,r,s_n}(k),
\]

where \( Δ x_{clk,r,s_n} ≜ \begin{bmatrix} c[Δ t_{r,s_n}, c[Δ t_{r,s_n}]]^\top \end{bmatrix} \), \( c[Δ t_{r,s_n}] ≜ c[Δ t_{r,s_n} − δ t_{r,s_n}(k)] \) is the difference between the clock drift of the receiver and the transmitter, and \( w_{clk,r,s_n} ≜ \begin{bmatrix} w_{δ t_{r,s_n}}, w_{δ t_{r,s_n}} \end{bmatrix}^\top \) is the process noise, which is modeled as zero-mean white random sequence with covariance

\[
Q_{clk,r,s_n} = Q_{clk,r} + Q_{clk,s_n},
\]

where \( Q_{clk,r} \) is obtained from (2). Recall that \( Q_{clk,s_n} \) has the same form as \( Q_{clk,r} \), except that \( S_{δ t_{r,s_n}} \) and \( S_{δ t_{r,s_n}} \) are now replaced by the transmitter-specific spectra \( S_{δ t_{r,s_n}} \) and \( S_{δ t_{r,s_n}} \), respectively. It is shown in Appendix A that \( η_{r,s_n}(k) \) is a zero-mean white random sequence with variance \( σ^2_{η,r,s_n} \) and is obtained from the \( k \)-th element of the vector

\[
η_{r,s_n} = GF_1 w_1 + GF_2 w_2,
\]

where \( G, F_1, \) and \( F_2 \) are deterministic matrices and

\[
w_1 ≜ \begin{bmatrix} w_{δ t_{r,s_n}(0)}, \ldots, w_{δ t_{r,s_n}(K - 2)} \end{bmatrix}^\top,
\]

\[
w_2 ≜ \begin{bmatrix} w_{δ t_{r,s_n}(0)}, \ldots, w_{δ t_{r,s_n}(K - 3)} \end{bmatrix}^\top,
\]

where \( K \) is the number of total processed samples. The derivation of \( σ^2_{η,r,s_n} \), \( G, F_1, \) and \( F_2 \) are presented in Appendix A. Fig. 2 shows \( σ_{η,r,s_n} \) for a transmitter equipped with a high-quality oven-controlled crystal oscillator (OCXO) and four different receivers: (i) a receiver equipped with a high-quality OCXO, (ii) a receiver equipped with a typical OCXO, a receiver equipped with a typical temperature-compensated crystal oscillator (TCXO), and (iv) a receiver equipped with a worst TCXO. From Fig. 2, one can see the conservativeness of the mismatch \( η_{r,s_n}(k) \) due to different clock types. In particular, the “inflation” in the calculated standard deviation of \( δ t_{r,s_n}(k) \) compared to the standard deviation of \( w_{δ t_{r,s_n}}(k) \) range between about one centimeter for (a) and about 2 m for (d).

---

**Fig. 1.** Cellular pseudorange measurement error characterization. (a) Pseudorange measurement after subtracting the true range between the receiver and the transmitter, which is denoted as \( z_{r,s_n}(k) \). (b) Resulting measurement after subtracting the clock bias difference between the receiver and the transmitter, which is denoted as \( ˜z_{r,s_n}(k) \). (c) The acf and (d) psd of \( ˜z_{r,s_n}(k) \) with a sampling frequency of 100 Hz.

**Fig. 2.** The standard deviation \( \{σ_{η,r,s_n}(k)\}_{k=0}^{K-1} \) of the vector \( η_{r,s_n} \) for four receivers, equipped with different oscillators: (a) high-quality OCXO, (b) typical OCXO, (c) typical TCXO, and (d) worst TCXO. Here, \( K = 2,500 \) samples and \( T = 0.01 \) s.

**Table II**

<table>
<thead>
<tr>
<th>Environment</th>
<th>( \mu )</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open sky</td>
<td>( 1.14 \times 10^{-15} ) m</td>
<td>0.11 m</td>
</tr>
<tr>
<td>Urban and semi-urban</td>
<td>( -1.24 \times 10^{-16} ) m</td>
<td>2.74 m</td>
</tr>
<tr>
<td>Deep urban</td>
<td>( -1.05 \times 10^{-15} ) m</td>
<td>5.42 m</td>
</tr>
</tbody>
</table>

It is worth comparing the values in Table II to GPS measurement’s user range error (URE), which are shown to have a standard deviation of around 6 m [53].

**B. Outlier Characterization**

This subsection characterizes two sources of pseudorange measurement outliers: (i) bias due to LOS signal blockage where the non-LOS (NLOS) peak dominates and (ii) bias due to short multipath delay where the LOS peak and multipath peeks are indiscernible.
in the estimated position [60]. Fig. 4 shows an instance of the estimated channel impulse response (CIR) from real LTE signals passing through a building. It can be seen that the power of the LOS component is significantly lower than that of the multipath component, which is at a path delay of approximately 440 m. This will introduce a large unmodeled bias in the pseudorange measurement. The objective of the outlier detection and exclusion approach discussed in Subsection III-C is to autonomously detect the presence of such bias and subsequently exclude the biased pseudorange measurement from the EKF calculations.

2) Bias Due to Short Multipath Delays: In the case of short multipath delays, i.e., the LOS peak is indiscernible from other multipath peaks, biases will be induced in the pseudorange measurement. This bias is characterized next. The true CIR for the $i$-th LTE symbol is given by

$$h_i(\tau) = \sum_{l=0}^{L-1} a_i(l) \delta(\tau - \tau_i(l)),$$

where $L$ is the number of path delays, $a_i$ corresponds to the complex-valued amplitude, and $\tau_i$ is the corresponding path delay. Note that path delays greater than the inverse of the signal bandwidth are excluded from the CIR since their effect will be negligible on the LOS component. The CIR can be used to measure the effect of multipath interference on the receiver’s delay-locked loop (DLL), denoted $\chi_n \triangleq \chi_{1,n}(t) +$
\[ \chi_2,n(i) = A \sum_{m=0}^{M-1} \sum_{l=1}^{L-1} a_i(l) e^{-j2\pi(m/M)(\tau_i(l)/T_s + \bar{\epsilon}_\theta)} \]
\[ \chi_1,n(i) = A \left| \frac{\sum_{m=0}^{M-1} \sum_{l=1}^{L-1} a_i(l) e^{-j2\pi(m/M)(\tau_i(l)/T_s + \bar{\epsilon}_\theta)}}{A} \right|^2 \]

where \( T_s \) is the subcarrier interval, \( \xi \) is the DLL correlator spacing, \( M \) is the number of subcarrier symbols in the LTE pilot signal, \( \text{Re} \{ \cdot \} \) denotes the real part, \( A \) is the signal power due to antenna gain and implementation loss, and \( \bar{\epsilon}_\theta \) is the normalized symbol timing error [61], which is set to zero to pilot signal, \( M \).

\[ \chi_2,n(i) = 2A \text{Re} \left\{ \left[ \sum_{m=0}^{M-1} e^{-j2\pi(m/M)(\bar{\epsilon}_\theta - \xi)} \right] \times \left[ \sum_{m'=0}^{M-1} \sum_{l=1}^{L-1} a_i^*(l) e^{j2\pi(m'/M)(\tau_i(l)/T_s + \bar{\epsilon}_\theta - \xi)} \right] \right\} \]
\[ -2A \text{Re} \left\{ \left[ \sum_{m=0}^{M-1} e^{-j2\pi(m/M)(\bar{\epsilon}_\theta + \xi)} \right] \times \left[ \sum_{m'=0}^{M-1} \sum_{l=1}^{L-1} a_i^*(l) e^{j2\pi(m'/M)(\tau_i(l)/T_s + \bar{\epsilon}_\theta + \xi)} \right] \right\} \]

where \( \kappa \equiv \frac{4\pi A \cos \left( \frac{\pi n}{M} \right)}{M \left| \sin \left( \frac{\pi n}{M} \right) \right|^3} \),

with a known distribution. Weighted least squares (WLS)-based filters, which use a weighted sum squared error (WSSE)-based test statistic have been demonstrated to result in an acceptable performance [62]. However, in [63], it was shown that using the normalized innovation squared (NIS)-based test statistic in an EKF-based framework was more robust against outliers compared to the WSSE-based test statistic in specific applications (e.g., railway navigation). In this paper, the NIS is used for generating the test statistic. The NIS-based test is formulated next.

Under outlier-free, normal operation, the innovation vector \( \nu_0(k+1) \) and its associated innovation covariance \( S(k+1) \) are given by

\[ \nu_0(k+1) \triangleq z(k+1) - \hat{z}(k+1|j), \]
\[ \approx \mathbf{H}(k+1)\hat{x}(k+1|j) + \nu'(k+1), \]
\[ S(k+1) \triangleq \mathbf{H}(k+1) \mathbf{P}(k+1|j) \mathbf{H}^T(k+1) + \mathbf{R}, \]

where \( \nu'(k+1) \triangleq [\nu'_{n1}, \ldots, \nu'_{nN}]^T \). Whenever the \( n \)-th measurement experiences a bias, the biased innovation vector \( \nu(k+1) \) may be expressed as

\[ \nu(k+1) = \nu_0(k+1) + \mathbf{u}_n(k+1), \]

where the vector \( \mathbf{u}_n(k+1) \triangleq [0, \ldots, 0, f_n(k+1), 0, \ldots, 0]^T \) is the bias vector that results when a bias of magnitude \( f_n \) is present in the pseudorange measurement drawn from the outlier cellular tower. Note that \( \nu(k+1) \) has the same covariance as \( \nu_0(k+1) \).

Denote \( \nu(k+1) \) as the innovation vector evaluated by the EKF. Subsequently, the following hypotheses may be posed:

- Null hypothesis:
  \[ H_0 : \nu(k+1) = \nu_0(k+1). \]
• Alternative hypothesis:

\[ H_1 : \ \nu(k+1) = \nu(k+1). \]

The hypothesis test relies on the NIS-based test statistic, which is defined as

\[ \varphi(k+1) \triangleq \nu^T(k+1)S^{-1}(k+1)\nu(k+1). \]

It is important to note that the NIS-based test statistics follows a chi-squared distribution under \( H_0 \) (outlier-free operation) and a non-central chi-squared distribution under \( H_1 \) (in the presence of outliers) [62]. Both distributions under \( H_0 \) and \( H_1 \) have the same degrees of freedom \( d = N_s \). The non-centrality parameter under \( H_1 \) (in the presence of outliers) is given by

\[ \lambda(k+1) = u_n^T(k+1)S^{-1}(k+1)u_n(k+1). \]

Outlier detection is achieved by comparing the test statistic against a detection threshold \( T_h \), namely

\[ \varphi(k+1) \leq T_h : H_0 \text{ is accepted (no outliers present)}, \]

\[ \varphi(k+1) > T_h : H_1 \text{ is accepted (outlier present)}, \]

where a Neyman-Pearson approach is taken to obtain \( T_h \) given a desired probability of false alarm \( P_{FA} \) according to

\[ P_{FA} = \int_{T_h}^{\infty} f_{\chi^2_d}(\tau)d\tau, \quad (23) \]

where \( f_{\chi^2_d} \) represents the pdf of the test statistic under \( H_0 \), specifically chi-squared distribution with \( d \) degrees of freedom.

Once a desired \( P_{FA} \) is fixed, \( T_h \) can be evaluated numerically from (23) or via chi-squared cumulative density function (cdf) table. Once an outlier is detected, its measurement is excluded from the measurement set and the EKF measurement update is subsequently performed without this measurement. The autonomous measurement outlier detection and exclusion algorithm is summarized in Algorithm 1.

### IV. Simulation Results

A simulation environment was set up to evaluate the performance of the proposed method described in Section III. The simulation settings are given in Table III. A ground vehicle was assumed to navigate in an urban environment (downtown Riverside, California) comprising six transmitters with known locations. The location of the transmitters were chosen according to real LTE towers in that environment. Over the course of this simulation, the vehicle traversed a 6 km trajectory, which included straight segments and turns. The LTE towers were assumed to be equipped with OCXOs, while the vehicle-mounted receiver was assumed to be equipped with a TCXO. The IMU’s rotational velocity and linear acceleration measurements were simulated at \( T = 0.01 \) s. The IMU’s measurement noise and time evolution of the IMU’s biases were simulated using parameters given in Table III. In Table III, \( \{ r_{sn} \}_{n=1}^{N_s} \) are expressed in an East, North, Up (ENU) coordinate frame. Fig. 7 illustrates the simulation environment layout, the trajectory traversed by the ground vehicle, and the position of cellular towers.

The model presented in Subsection III-B2 was used to simulate the effect of short multipath delays on the measurements.

The channel was assumed to be Rician with an extended Vehicular A (EVA) power delay profile. Therefore, \( a_0 \) was set to 1 and \( \{ a_i \}_{i=1}^{L-1} \) were assumed to be zero-mean Gaussian random variable with variance \( \sigma_{a_i}^2 \), whose values are given in Table III. The details of the EVA channel model are presented in [64]. The simulated induced multipath biases \( \{ b_n \}_{n=1}^{N_s} \) obtained from (21) for all six transmitters are shown in Fig. 6. These biases were added into the pseudorange measurements drawn from the six cellular towers.

To simulate the effect of a bias due to LOS signal blockage, discussed in Subsection III-B1, a step bias with a magnitude of 25 m was introduced into the pseudorange drawn from cellular tower 3, starting at \( t = 180 \) s.

The autonomous measurement outlier detection and exclusion approach formulated in Subsection III-C was executed on the simulated data. Fig. 8 shows the outlier detection test obtained by comparing the test statistic \( \varphi \) against the detection threshold \( T_h \) for (a) in the outlier-free case and (b) in the presence of the outlier. Fig. 8(c)-(d) show the test statistic \( \varphi \) distribution (as calculated from simulated data) for the outlier-free case and in the presence of the outlier, respectively. Overlaid on these plots are the analytical distributions discussed in Subsection III-C: (i) central chi-squared with \( d = 6 \) and (ii)

---

**Algorithm 1: Autonomous measurement outlier detection and exclusion**

<table>
<thead>
<tr>
<th>Step</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Set ( i = 1 )</td>
</tr>
<tr>
<td>2</td>
<td>Set ( \varphi(k+1) = \nu^T(k+1)S^{-1}(k+1)\nu(k+1) )</td>
</tr>
<tr>
<td>3</td>
<td>If ( \varphi(k+1) &gt; T_h )</td>
</tr>
<tr>
<td>4</td>
<td>Exclude ( z_s ) from ( { z_{sn}(k+1) }_{n=1}^{N_s} )</td>
</tr>
<tr>
<td>5</td>
<td>Set ( \varphi(k+1) = \nu^T(k+1)S^{-1}(k+1)\nu(k+1) )</td>
</tr>
<tr>
<td>6</td>
<td>If ( \varphi(k+1) \leq T_h )</td>
</tr>
<tr>
<td>7</td>
<td>Add ( z_s ) to ( { z_{sn}(k+1) }_{n=1}^{N_s-1} )</td>
</tr>
<tr>
<td>8</td>
<td>Set ( i = i + 1 )</td>
</tr>
<tr>
<td>9</td>
<td>If ( i \leq N_s )</td>
</tr>
<tr>
<td>10</td>
<td>Go to Step 4</td>
</tr>
<tr>
<td>11</td>
<td>Else</td>
</tr>
<tr>
<td>12</td>
<td>RAIM is not available: there is an outlier, but</td>
</tr>
<tr>
<td>13</td>
<td>it cannot be detected</td>
</tr>
<tr>
<td>14</td>
<td>Exit Algorithm</td>
</tr>
<tr>
<td>15</td>
<td>End if</td>
</tr>
<tr>
<td>16</td>
<td>Else</td>
</tr>
<tr>
<td>17</td>
<td>( i )-th measurement is declared as an outlier</td>
</tr>
<tr>
<td>18</td>
<td>Feed navigation block with</td>
</tr>
<tr>
<td>19</td>
<td>( { z_{sn}(k+1) }_{n=1}^{N_s-1} )</td>
</tr>
<tr>
<td>20</td>
<td>Exit Algorithm</td>
</tr>
<tr>
<td>21</td>
<td>End if</td>
</tr>
<tr>
<td>22</td>
<td>Else</td>
</tr>
<tr>
<td>23</td>
<td>No outlier is detected</td>
</tr>
<tr>
<td>24</td>
<td>Feed navigation block with ( { z_{sn}(k+1) }_{n=1}^{N_s} )</td>
</tr>
<tr>
<td>25</td>
<td>Exit Algorithm</td>
</tr>
<tr>
<td>26</td>
<td>End if</td>
</tr>
</tbody>
</table>
non-central chi-squared with $d = 6$ and $\lambda = 23.71$.

Fig. 9 shows the position estimation error trajectories and corresponding $\pm 3\sigma$ for (a)-(b): outlier-free, (c)-(d): in the presence of outlier, and (e)-(f): in the presence of outlier followed by the proposed outlier detection and exclusion approach.

It can be seen from 9 (c)-(d) and (e)-(f) that not only the detection and exclusion algorithm reduces the position estimation error, but also it results in estimation errors that are consistent with $\pm 3\sigma$ bounds. This has important implications on the integrity of the navigation solution. Table IV summarizes the 2-D position RMSE for the three scenarios.

V. EXPERIMENTAL RESULTS

A field test with real LTE signals was performed to validate the efficacy and accuracy of the proposed outlier detection and exclusion approach. In this section, the experimental setup and experimental scenario description are first presented. Then, the corresponding results are shown.

A. Experimental Hardware Setup and Scenario Description

A vehicle was equipped with following hardware and software setup:

- A Septentrio AsteRx-i V integrated GNSS-IMU whose $x$-axis points toward the front of the vehicle, $z$-axis points upward, and $y$-axis points to the right side of the car. AsteRx-i V is equipped with a dual-antenna,
shown are analytical distributions of $P_\phi$; where in (c) a central chi-squared distribution with $d = 6$ and in (d) a non-central chi-squared with $d = 6$ and $\lambda = 23.71$ are plotted.

Fig. 8. Top: Test statistic $\varphi$ under (a) outlier-free condition and (b) in the presence of an outlier. Also, shown is the threshold $T_{\varphi}$, calculated based on a desired $P_\varphi$. Bottom: Test statistic $\varphi$ distribution calculated from simulated data under (c) outlier-free condition and (d) in the presence of an outlier. Also shown are analytical distributions of $\varphi$; where in (c) a central chi-squared distribution with $d = 6$ and in (d) a non-central chi-squared with $d = 6$ and $\lambda = 23.71$ are plotted.

### TABLE IV

**Comparison between navigation solution performance**

<table>
<thead>
<tr>
<th>Scenario</th>
<th>2-D RMSE [m]</th>
<th>2-D Max. error [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outlier-free</td>
<td>0.39</td>
<td>1.52</td>
</tr>
<tr>
<td>In the presence of outlier without detection and exclusion</td>
<td>3.67</td>
<td>16.26</td>
</tr>
<tr>
<td>In the presence of outlier after detection and exclusion</td>
<td>0.42</td>
<td>3.40</td>
</tr>
<tr>
<td>Improvement</td>
<td>88.56%</td>
<td>79.09%</td>
</tr>
</tbody>
</table>

multi-frequency GNSS receiver and a VectorNav VN-100 micro-electromechanical system (MEMS) IMU. The raw measurements from the IMU was used for the EKF time update step described in Subsection II-B at a rate of 100 Hz.

- Septentrio’s post-processing software development kit (PP-SDK) was used to process carrier phase observables collected by the AsteRx-i V and by a nearby differential GPS base station to obtain a carrier phase-based navigation solution. This integrated GNSS-IMU real-time kinematic (RTK) system [65] was used to produce the ground truth results with which the proposed navigation framework was compared.

- Two cellular antennas to acquire and track signals from nearby cellular LTE towers. The LTE antennas used for the experiment were consumer-grade 800/1900 MHz cellular antennas [66]. The signals were simultaneously down-mixed and synchronously sampled at 10 mega-samples per second (MSPS) via a National Instruments (NI) dual-channel universal software radio peripheral (USRP)–2954R, driven by a GPSDO [67]. These samples were then processed by MATRIX SDR [21], [27], [28], developed by the Autonomous Systems Perception, Intelligence, and Navigation (ASPIN) Laboratory at the University of California, Irvine.

A stationary base receiver with the knowledge of its own location was also deployed in the same cellular environment of the navigating vehicle and was equipped with the following hardware setup:

- Two consumer-grade 800/1900 MHz cellular antennas.
- Two NI USRP–2920 which were synchronized via a multiple-input and multiple-output (MIMO) cable.

The clock bias and drift process noise power spectral densities of the receiver were set to be $1.3 \times 10^{-22}$ s and $7.89 \times 10^{-25}$ 1/s respectively, since the 2954R USRP was equipped with OCXO. The receiver was tuned to carrier frequencies of 1955 MHz and 739 MHz, which are channels allocated for U.S. cellular provider AT&T. Samples of the received signals were stored for off-line post-processing. The SDR developed in [27] was used to produce LTE pseudoranges.

Fig. 10 illustrates the experimental hardware setup, experimental environment, and traversed trajectory along with the position of the base and cellular towers. The initial estimates of the vehicle’s orientation $I^{-1}_3 \tilde{q}(0) - 1$, position $G \tilde{r}_r(0) - 1$, and velocity $G \tilde{v}_r(0) - 1$, and their covariances were initialized using the output of the GNSS-IMU system. The initial estimates of the gyroscope’s and accelerometer’s biases; $b_q(0)$ and $b_a(0)$, respectively; were obtained by averaging 5 seconds of IMU measurements at a sampling period of $T = 0.01$ seconds, while the vehicle was stationary. Since the vehicle had initial access to GNSS signals, the initial value of the vehicle-mounted receiver’s clock error states $\hat{x}_{clk,r}(0) - 1$ was obtained. The initial uncertainties associated with these state estimates were set to $P_q(0) - 1 = (1 \times 10^{-20}) I_{3 \times 3}$, $P_{G\tilde{r}_r}(0) - 1 = \text{diag}(10, 10, 0)$, $P_{G\tilde{v}_r}(0) - 1 = \text{diag}(0.5, 0.5, 0)$, $P_{b_q}(0, 0) - 1 = (3.75 \times 10^{-9}) I_{3 \times 3}$, $P_{b_a}(0, 0) - 1 = (9.6 \times 10^{-5}) I_{3 \times 3}$, and $P_{x_{clk,r}(0, 0) - 1} = \text{diag}(3, 0, 3)$, where diag($\cdot$) denotes a diagonal matrix.

The probability of false alarm for the outlier detection test was set to $P_{FA} = 0.005$. The measurement noise variances were calculated empirically while the vehicle had access to GNSS signals according to

$$
\sigma^2_{\tilde{v},s_n} \approx \frac{1}{k_{cutoff} - 1} \sum_{k=0}^{k_{cutoff} - 1} \tilde{v}_{s_n}^2(k),
$$

where $k_{cutoff}$ is the time-step at which GNSS signals were cut off and

$$
\tilde{v}_{s_n}^2(k) \triangleq \tilde{z}_{s_n}(k) - \left\| G_{r_r,GNSS}(k) - r_{s_n}(k) \right\|_2^2 - c \left\| \tilde{\delta t}_{r_r,GNSS}(k) - \tilde{t}_{s_n}(k) \right\|_2,
$$

where $G_{r_r,GNSS}$ and $\delta t_{r_r,GNSS}$ are the vehicle-mounted receiver’s position and clock bias estimates obtained using GNSS signals and $\tilde{t}_{s_n}$ is the clock bias estimate produced by the base.

Over the course of the experiment, the receiver was listening to 5 LTE towers with the characteristics shown in Table V. The receiver traversed a trajectory of 1500 m over 135 s.
Fig. 9. Position estimation errors and corresponding ±3σ for three simulation scenarios; (a)-(b): outlier-free, (c)-(d): in the presence of outlier, and (e)-(f): in the presence of outlier operation followed by the proposed outlier detection and exclusion technique.

B. Experimental Results

The outlier detection and exclusion approach was applied throughout the vehicle’s traversed trajectory. Fig. 11 shows the outlier detection test, which compares the test statistic ϕ against the detection threshold $T_h$. It can be seen that at $t = 100$ s, the threshold is exceeded; therefore, the test is not declared successful (see the red circle in Fig. 11). This implies that at least one of the measurements was detected as an outlier.
TABLE V
LTE Towers’ Characteristics

<table>
<thead>
<tr>
<th>eNodeB</th>
<th>Carrier frequency (MHz)</th>
<th>Cell ID</th>
<th>Bandwidth (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1955</td>
<td>216</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>739</td>
<td>319</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>739</td>
<td>288</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>739</td>
<td>151</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>739</td>
<td>232</td>
<td>10</td>
</tr>
</tbody>
</table>

Fig. 11. The resulting outlier detection test which compares the test statistic \( \varphi \) against the detection threshold \( T_h \).

and its contribution to the test statistic was significant enough for the detection algorithm to trigger.

In this experiment, the outlier exclusion technique indicated that the outlier measurement was the pseudorange drawn from the second cellular tower. Fig. 12 (a)–(b) shows the resulting position estimation error and corresponding ±3σ bounds with and without using the proposed outlier exclusion. As can be seen, outlier exclusion results in a significant reduction in the \( x \)- and \( y \)-position estimation error.

The pdf of the test statistic as calculated from the experimental data is plotted in Fig. 13 for the (a) outlier-free case (i.e., after the outlier was detected and excluded) and (b) in the presence of outlier (i.e., before it was detected and excluded). Overlaid on each pdf is an analytical pdf of (a) central chi-squared distribution with \( d = 5 \) and (b) non-central chi-squared distribution with \( d = 5 \) and \( \lambda = 12 \). It can be seen that the empirical pdfs follow the analytical pdfs.

Table VI compares the navigation performance of the proposed framework versus that of the navigation solution without outlier exclusion. Similar to the simulation results, it can be noted from Fig. 12 and Table VI, not only the outlier detection and exclusion reduced the estimation error, it also resulted in estimation error consistent with the ±3σ bounds.

VI. Conclusion

In this paper, a framework for the ground vehicle outlier detection and exclusion in GNSS–denied environment was developed. To this end, an EKF-based navigation framework was proposed which used IMU data and pseudoranges extracted from ambient cellular LTE towers. This paper analyzed the LTE pseudorange measurements’ characterization using the recorded data in different environments: open sky, urban, and deep urban. Also, it characterized two main sources of
LTE pseudorange measurements’ outliers: (i) LOS blockage or high attenuation and (ii) short multipath delays. Next, an autonomous outlier detection and exclusion approach was formulated and evaluated using realistic simulation tests and an experimental test. Experimental results over a total traversed trajectory of 1500 m validated the efficacy of the proposed framework and also showed that the proposed outlier detection and exclusion technique reduces the position 2-D RMSE by 41.5%. While this paper considered a classical RAIM, more sophisticated RAIM algorithms, such as advanced RAIM (ARAIM), could be investigated in future work in an attempt to (i) account for the multi-outliers conditions and (ii) develop a multi-constellation-based approach that combines the LTE signals and GNSS signals to enhance the safety of the ground driving.

APPENDIX A
DERIVATION OF EQUATION (18)

In this Appendix, the mismatch between the time evolution of true clock bias and its first-order polynomial approximation is analyzed and the error component due to this approximation, i.e., $\eta_{r,s_n}$, is characterized. The first step is to estimate the constants initial clock bias $c\delta t_{r,s_n,0}$ and drift $c\dot{\delta}t_{r,s_n,0}$ and find the first-order polynomial approximation over all observations up to time-step $K-1$. It is important to note that the measurement noise will not be considered here as the mismatch between true clock bias and its first-order polynomial approximation is only affected by the process noise. Hence, the observation vector only includes the differences between the receiver’s and transmitter’s clock biases, i.e.,

$$y \triangleq [c\Delta \delta t_{r,s_n}(0), \ldots, c\Delta \delta t_{r,s_n}(K-1)^\top].$$

A first-order polynomial can be fit to the observations using a least-squares estimator, which minimizes the cost function $G$ according to

$$G = \|y - S \begin{bmatrix} c\dot{\delta}t_{r,s_n,0} \\ c\dot{\delta}t_{r,s_n,0} \end{bmatrix}\|^2,$$

where $S \triangleq \begin{bmatrix} 0 & T & \ldots & (K-1)T \\ 1 & 1 & \ldots & 1 \end{bmatrix}^\top$.

over all possible $c\dot{\delta}t_{r,s_n,0}$ and $c\dot{\delta}t_{r,s_n,0}$.

The least-squares-based estimate of $c\dot{\delta}t_{r,s_n,0}$ and $c\dot{\delta}t_{r,s_n,0}$ is given by

$$\begin{bmatrix} \hat{c}\dot{\delta}t_{r,s_n,0} \\ \hat{c}\dot{\delta}t_{r,s_n,0} \end{bmatrix} = (S^\top S)^{-1}S^\top y,$$

and the mismatch error vector can be expressed as

$$\eta_{r,s_n} = y - S \begin{bmatrix} c\dot{\delta}t_{r,s_n,0} \\ c\dot{\delta}t_{r,s_n,0} \end{bmatrix} = Gy.$$

where $G = [I - S(S^\top S)^{-1}S^\top]$. Using (17), $c\Delta \delta t_{r,s_n}(k)$ can be expressed in the recursive form according to

$$c\Delta \delta t_{r,s_n}(k) = c\Delta \delta t_{r,s_n}(0) + kTc\dot{\delta}t_{r,s_n}(0)$$

$$+ \sum_{i=1}^{k-1} w_{\delta t_{r,s_n}}(i) + \sum_{i=1}^{k-2} (k-i-1)w_{\delta t_{r,s_n}}(i).$$

Substituting (27) into (25) results in

$$y = S \begin{bmatrix} c\Delta \delta t_{r,s_n}(0) \\ c\Delta \delta t_{r,s_n}(0) \end{bmatrix} + F_1 w_1 + F_2 w_2,$$

where

$$F_1 \triangleq \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \ldots & 1 \end{bmatrix}^{(K) \times (K-1)},$$

$$F_2 \triangleq \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ T & 0 & \ldots & 0 \\ 2T & T & \ldots & 0 \end{bmatrix}^{(K-2) \times (K-3) \times T} \ldots T^{(K-2)}$$

and recall

$$w_1 \triangleq \begin{bmatrix} w_{\delta t_{r,s_n}}(0), \ldots, w_{\delta t_{r,s_n}}(K-2) \end{bmatrix}^\top,$$

$$w_2 \triangleq \begin{bmatrix} w_{\delta t_{r,s_n}}(0), \ldots, w_{\delta t_{r,s_n}}(K-3) \end{bmatrix}^\top.$$

Substituting (28) into (26) yields

$$\eta_{r,s_n} = GS \begin{bmatrix} c\Delta \delta t_{r,s_n}(0) \\ c\Delta \delta t_{r,s_n}(0) \end{bmatrix} + GF_1 w_1 + GF_2 w_2,$$

Since $GS = 0$, (29) can be expressed as

$$\eta_{r,s_n} = GF_1 w_1 + GF_2 w_2.$$

This shows that each element of $\eta_{r,s_n}$ is a linear combination of $2K-3$ zero-mean white noise, therefore, $\eta_{r,s_n}$ is a zero-mean white random vector with a covariance matrix given by

$$E\{\eta_{r,s_n} \eta_{r,s_n}^\top\} = \begin{bmatrix} GF_1 w_1 F_1^\top G^\top + GF_2 w_2 F_2^\top G^\top \\ + GF_1 w_1 F_2^\top G^\top + GF_2 w_2 F_1^\top G^\top \end{bmatrix} + q_{11} GF_1 F_1^\top G^\top + q_{22} GF_2 F_2^\top G^\top$$

$$+ q_{12} G \begin{bmatrix} F_1 H_q F_2^\top + F_2 H_q F_1^\top \end{bmatrix} G^\top,$$

where $H_q = [I_{(K-2) \times (K-2)} \ 0_{(K-2) \times 1}]^\top$ and $q_{ij} = ij$-th element of the matrix $Q_{clk,r,s_n}$.  

**Remark:** By extending the expression of matrix $G$ as a function of $k$ and $T$, it can be shown that $G$ is a bisymmetric matrix. The covariance matrix of $\eta_{r,s_n}$ consists of three terms shown in (30). By constructing $F_1 F_1^\top$, $F_2 F_2^\top$, and $F_1 H_q F_2^\top + F_2 H_q F_1^\top$, it can be shown that each of the three terms of the covariance matrix are also bisymmetric. As a result, the diagonal elements of the covariance matrix of $\eta_{r,s_n}$ are symmetric, as it is evident from Fig. 2.

**ACKNOWLEDGMENT**

The authors would like to thank Joe Khalife and Kimia Shamaei for her help with the data collection.
REFERENCES


