

Subspace-based Adaptive GMM Error Modeling for Fault-Aware Pseudorange-based Positioning in Urban Canyons

Penggao Yan, Xiao Xia, Michele Brizzi, Weisong Wen, Li-Ta Hsu *Senior Member, IEEE*

Abstract—Global navigation satellite system (GNSS) positioning is essential for achieving absolute vehicular positioning in urban scenarios; however, it suffers from limited measurement redundancy and substantial faults caused by complex urban environments. In this work, we propose the subspace-based adaptive error modeling and fault detection and exclusion (FDE) method for pseudorange-based GNSS positioning in urban canyons, which integrates the adaptive error modeling into the FDE process and the positioning-solving process. Notably, we divide the pseudorange measurement space into subspaces regarding elevation angle and carrier-to-noise ratio (C/N0), each of which maintains a Gaussian mixture model (GMM) to adaptively characterize measurement error profiles. Results show that the proposed method has the ability to detect environmental changes. In addition, the proposed method outperforms the conventional FDE method with Gaussian assumptions, reducing the mean positioning error by 16% and 9% in slightly and medium urbanized datasets, respectively. The impacts of step size (elevation angle and C/N0) and time window of the proposed method are discussed through controlled experiments.

Index Terms—Adaptive error modeling, fault detection, global navigation satellite system, Gaussian mixture model, urban canyons

I. INTRODUCTION

GLOBAL navigation satellite system (GNSS) positioning plays a crucial role in enabling intelligent vehicles to navigate and operate effectively in urban scenarios by providing absolute positioning information [1]. By leveraging the information of absolute positioning, intelligent vehicles can precisely determine their position within the urban environment, enabling them to make informed decisions, such as lateral control [2], collision avoidance [3], [4], and route planning [5], [6]. Although the map-based LiDAR positioning method can also provide absolute positioning information, it comes with certain drawbacks, such as the high cost associated with map maintenance [7] and the vulnerability of LiDAR to adverse weather conditions [8], [9]. Therefore, GNSS stands out as a cost-effective and indispensable solution for the absolute positioning of intelligent vehicles in urban scenarios.

Penggao Yan, Xiao Xia, Weisong Wen, and Li-Ta Hsu are with the Department of Aeronautical and Aviation Engineering, Hong Kong Polytechnic University, Hong Kong (e-mail: peng-gao.yan@connect.polyu.hk; xxiao.xia@connect.polyu.hk; welson.wen@polyu.edu.hk; lt.hsu@polyu.edu.hk).

Michele Brizzi is with the Department of Industrial, Electronic and Mechanical Engineering, Roma Tre University, Italy (e-mail: michele.brizzi@uniroma3.it).

The corresponding author: Li-Ta Hsu (e-mail: lt.hsu@polyu.edu.hk).

However, the complex urban environments pose substantial challenges for GNSS positioning [10]. The GNSS signals can be affected by various environmental objects, such as high buildings and trees, which results in the reflection, diffusion, or blockage of the GNSS signals, consequently reducing the reliability of GNSS measurement as well as introducing substantial measurement faults [10], [11].

Fault detection and exclusion methods with Gaussian error modeling: One dominant solution to handle faults in GNSS measurements is employing fault detection and exclusion (FDE), which is the technology to check the occurrence of faults in the system and to determine the time when faults occur [12]. In this paper, we define faults as unmodelled measurement errors that substantially compromise the positioning solution, such as non-line-of-sight (NLOS) errors [13]–[17]. Those measurement errors that have relatively smaller impacts and are induced by environments are regarded as nominal measurement errors. Indeed, there is some intersection between the nominal measurement errors and the faults, which makes it difficult to distinguish the difference in a sole GNSS positioning system. Therefore, the definition of faults and nominal measurement error emphasizes the impacts of measurements on the positioning solution. The fundamental principle of FDE is to identify measurements that deviate from the system's expected behavior, where abnormal measurements usually show a considerable inconsistency with normal measurements [18]. Researchers have been utilizing this inconsistency to develop various FDE methods, such as parity space [19], chi-squared test [16], [20], and solution separation [13], [21], which are further intergrated in receiver autonomous integrity monitoring (RAIM) [14], [15] and advanced RAIM (ARAIM) [13], [16] algorithms for integrity monitoring purpose. These methods usually assume that GNSS measurement errors are Gaussian-distributed. Indeed, this assumption brings several benefits, such as requiring fewer parameters (only two) in describing the distribution and two uncorrelated Gaussian variables are mutually independent. The latter property is actually the theoretical basis for developing the parity space and Chi-squared method [19]. However, many studies have exemplified that GNSS measurement errors have non-Gaussian properties [22], [23]. The unrealistic Gaussian assumptions can result in degraded fault detection rates in real-world applications [24], limiting the reliability and effectiveness of preventing systems from faults.

Fault detection and exclusion methods with non-Gaussian error modeling: To remedy these non-Gaussian issues, re-

searchers explore the possibility of incorporating non-Gaussian error modeling in developing fault detection algorithms. Two distinct approaches emerge, including the particle filter (PF) [25]–[27] and the Gaussian sum filter (GSF) [28], [29] approaches. In the PF approach, the measurement set is typically partitioned into multiple subsets. Within each subset, a PF with non-Gaussian error modeling is employed to estimate the state and its likelihood. Then a cumulative log-likelihood ratio (LLR) test is constructed by using the likelihood produced by each auxiliary PF and the main PF [26]. However, the PF approach necessitates a significant computational burden, which makes it difficult to fulfill the real-time requirements for vehicular positioning. In the GSF approach, the Gaussian mixture model (GMM) is employed to model measurement noise, and several parallel Kalman filters are utilized to address each Gaussian component. Residuals generated from each subfilter are summed up according to the mixture weight of GMM, based on which a chi-squared test is conducted to detect potential faults [28].

Static error modeling is inadequate in urban scenarios: It is easy to find that the Gaussian and GSF approaches involve modeling nominal measurement error, i.e., the measurement error in fault-free conditions. In the aviation field, the exact profile of GNSS measurement errors in open-sky conditions has been identified based on a large number of observations [20], [30]. However, the situation becomes contradictory in urban scenarios. Due to the impacts of various environmental objects, the GNSS measurement is distorted and its error profile changes with environment [10], [11]. To tackle these issues, goGPS has been proposed [31]. GoGPS models the GNSS measurement error as Gaussian distribution, whose variance is identified as a function of elevation angle and carrier-to-noise ratio (C/N0) [31]. However, this method has an underlying assumption that the uncertainty of GNSS measurements with the same elevation angles and C/N0 values remains constant across different urban scenarios; in other words, the error modeling process is static and primarily based on historical data, which may not capture the dynamic and diverse characteristics of urban environments. Although users can choose different weighting schemes by adjusting parameters in goGPS to accommodate the current environment, a substantial amount of data must be collected to determine the optimal parameters, which is impractical for real-time GNSS vehicular positioning systems in dynamic scenarios [32], [33].

Adaptive error modeling can be substantially affected by faulty measurements: Recently, Pfeifer and Protzel demonstrated a promising approach to solve the issues in the static error modeling. They proposed the concept of multimodal and adaptive error modeling [34] and embedded it in the robust estimation architecture [35], aiming to estimate positioning states and parameters of the measurement error distribution simultaneously. In their work, the GNSS measurement error model is represented by a GMM, and its parameters are updated by fitting a GMM model to historical measurement residuals. This method can adjust the weights of measurements based on the updated error profile, making it possible to adapt to environmental changes. However, this method did not consider the impacts of faulty measurements on identifying

the measurement error distribution with the residual data. Specifically, residuals are obtained by differentiating the raw measurements from predicted ones based on the measurement model and the estimated state, both of which could be substantially affected by faulty measurements. Therefore, a countermeasure is needed to reduce the impacts of faulty measurements on the quality of residual data before error modeling. In addition, this method assumes that GNSS measurements obtained from different satellite-user geometry have the same error profile, which is not practical even in open-sky conditions, let alone in urban scenarios.

To tackle these issues in the adaptive error modeling method, we propose the subspace-based adaptive GMM error modeling and FDE method by combining the principles of FDE with the adaptive error modeling process, providing a practical solution for accurate GNSS positioning in urban scenarios. Specifically, we divide the GNSS pseudorange measurement space into a set of distinct subspaces based on elevation angle and carrier-to-noise ratio (C/N0), each of which is represented with a bin. At each epoch that receives new GNSS measurements, we employ the FDE technique to identify “healthy” measurements, which are utilized to compute the measurement residuals. Then, each residual is assigned to the best-matched bin according to its associated elevation angle and C/N0. At the same time, outdated residual data is removed based on a predefined window size. For each updated bin, we utilize the collected residuals to fit a GMM distribution, which serves as the error model for GNSS pseudorange measurements within that particular subspace. In the subsequent epoch, the updated error model is utilized to calculate the weighting matrix based on the law of total variance. This weighting matrix is then adopted by both the FDE and position solver to enhance their performance.

The proposed method is evaluated on two urban datasets, including a slightly urban dataset and a medium urban dataset. The evolution of error models is compared with the change pattern of surroundings, where a strong relationship between them are found. In addition, results show that the surrounding change at a short period can only affect the error models with specific elevation angles, indicating that the partition of the measurement space is beneficial to establish a more accurate error model for GNSS pseudorange measurements. Furthermore, the proposed method is compared with the conventional FDE method with Gaussian assumption, robust estimation with Gaussian error modeling, and robust estimation with adaptive GMM error modeling, regarding the positioning performance. On both datasets, the proposed method demonstrates the best positioning performance, where a 16% and a 9% reduction of mean positioning error are achieved compared to the conventional FDE method with Gaussian assumption. Moreover, the superiority of the proposed method to the robust estimation with adaptive GMM error modeling reveals that the integration of the FDE process reduces the impact of environments on characterizing the error profile of GNSS measurements.

The contributions of this work are as three folds:

- 1) Propose a subspace-based adaptive GMM error modeling method for GNSS measurements in urban scenarios, which provides a cost-effective way to capture changes

in environments, serving as the basis for determining the weighting scheme in positioning solver and FDE.

- 2) Propose a practical framework for detecting and excluding faulty GNSS measurements based on the adaptive GMM error modeling, which strengthens the adaptive ability of GNSS positioning systems to changing environments, thereby enhancing GNSS positioning performance in urban canyons.
- 3) Experimentally demonstrates the effectiveness of the proposed framework in two distinct urban datasets. Furthermore, we discuss the impacts of step size (elevation angle and C/N0) and time window on the positioning performance of the proposed method, tentatively giving recommendations for practice in urban areas.

The rest of this article is organized as follows. Section II describes the conventional FDE method with Gaussian assumptions for GNSS positioning systems and points out its limitation in urban scenarios. Section III first introduces the feature of the GMM and proposes the subspace-based adaptive GMM error modeling and FDE method for GNSS positioning in urban areas. In Section IV, we examine the performance of the proposed method in error modeling, fault detection, and positioning on two urban datasets. In Section V, we discuss the impacts of step size (elevation angle and C/N0) and time window parameters in the proposed method. Section VI gives a summary.

II. FAULT DETECTION WITH GAUSSIAN ERROR MODELING

A. Weighted Least Square Estimation in GNSS

The GNSS pseudorange measurement model can be formalized as follows [1]:

$$\rho_i = \sqrt{(x_{s,i} - x)^2 + (y_{s,i} - y)^2 + (z_{s,i} - z)^2} + c\delta_r + \varepsilon_i, \quad (1)$$

where ρ_i is the i th pseudorange measurement, $[x_{s,i}, y_{s,i}, z_{s,i}]^T$ is the position of i th satellite, $[x, y, z]^T$ is the receiver position, δ_r is the receiver clock bias from single satellite constellation (GPS or BeiDou), $c = 3 \times 10^8$ m/s is the speed of light, and ε_i is the measurement error. In this work, the pseudorange measurement is corrected by applying the Differential Global Navigation Satellite System (DGNSS) technique, which removes common errors including satellite clock errors, ephemeris errors, ionospheric errors, and tropospheric errors in pseudorange measurement [1]. The reference station clock offset is also compensated. All processes are accomplished by using RTKLlib [36]. The right hand side (RHS) of (1) can be linearized by taking the first-order Taylor expansion at $\mathbf{x}_0 = [x_0, y_0, z_0, \delta_{r,0}]^T$ as follows:

$$\begin{aligned} \rho_i = & \rho_{i,0} - a_{i,1}(x - x_0) - a_{i,2}(y - y_0) - a_{i,3}(z - z_0) \\ & + c(\delta_r - \delta_{r,0}) + \varepsilon_i, \end{aligned} \quad (2)$$

where

$$\begin{aligned} \rho_{i,0} &= \sqrt{(x_{s,i} - x_0)^2 + (y_{s,i} - y_0)^2 + (z_{s,i} - z_0)^2} + c\delta_{r,0} \\ a_{i,1} &= \frac{x_{s,i} - x_0}{\sqrt{(x_{s,i} - x_0)^2 + (y_{s,i} - y_0)^2 + (z_{s,i} - z_0)^2}} \\ a_{i,2} &= \frac{y_{s,i} - y_0}{\sqrt{(x_{s,i} - x_0)^2 + (y_{s,i} - y_0)^2 + (z_{s,i} - z_0)^2}} \\ a_{i,3} &= \frac{z_{s,i} - z_0}{\sqrt{(x_{s,i} - x_0)^2 + (y_{s,i} - y_0)^2 + (z_{s,i} - z_0)^2}}. \end{aligned} \quad (3)$$

The matrix form of the linearized pseudorange measurement model with n measurements from two constellations (e.g., GPS and BeiDou) can be written as

$$\mathbf{y} = \mathbf{H}\Delta\mathbf{x} + \boldsymbol{\varepsilon}, \quad (4)$$

where

$$\begin{aligned} \mathbf{y} &= \begin{bmatrix} \rho_{1,0} - \rho_1 \\ \vdots \\ \rho_{k,0} - \rho_k \\ \rho_{k+1,0} - \rho_{k+1} \\ \vdots \\ \rho_{n,0} - \rho_n \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k,1} & a_{k,2} & a_{k,3} & 1 & 0 \\ a_{k+1,1} & a_{k+1,2} & a_{k+1,3} & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & 0 & 1 \end{bmatrix}, \\ \Delta\mathbf{x} &= \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \\ -c(\delta_r^G - \delta_{r,0}^G) \\ -c(\delta_r^B - \delta_{r,0}^B) \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_k \\ \varepsilon_{k+1} \\ \vdots \\ \varepsilon_n \end{bmatrix}, \end{aligned} \quad (5)$$

δ_r^G and δ_r^B are the receiver clock biases regarding the two constellations, respectively; and $\delta_{r,0}^G$ and $\delta_{r,0}^B$ are the corresponding linearized points. The exact forms of \mathbf{H} and $\Delta\mathbf{x}$ depend on the number of constellations and the ordering of the pseudorange measurements. (4) and (5) only give an example of a two-constellation GNSS system without loss of generality.

With n pseudorange measurements, the estimated receiver position $\hat{\mathbf{x}}$ can be solved by the weighted least square (WLS) method (in an iterative approach) as follows:

$$\begin{aligned} \Delta\hat{\mathbf{x}} &= (\mathbf{H}^T \mathbf{W} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W} \mathbf{y} \\ \hat{\mathbf{x}} &= \mathbf{x}_0 + \Delta\hat{\mathbf{x}}, \end{aligned} \quad (6)$$

where \mathbf{W} is the weighting matrix and usually takes the inverse of the covariance matrix of $\boldsymbol{\varepsilon}$. Here, $\hat{\mathbf{x}}$ stands for the real time differential (RTD) positioning result since its solution process uses the DGNSS technique for measurement error correction.

B. Fault Detection with Gaussian Assumption

The estimated pseudorange measurement error (measurement residual) $\hat{\boldsymbol{\varepsilon}}$ can be formalized by

$$\begin{aligned} \hat{\boldsymbol{\varepsilon}} &= \mathbf{y} - \mathbf{H}\Delta\hat{\mathbf{x}} \\ \hat{\boldsymbol{\varepsilon}} &= (\mathbf{I} - \mathbf{P})\mathbf{y}, \end{aligned} \quad (7)$$

where

$$\mathbf{P} = \mathbf{H} (\mathbf{H}^T \mathbf{W} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W}. \quad (8)$$

The weighted sum of the squared errors (WSSE) can be constructed as the test statistic [20]

$$\text{WSSE} = \hat{\boldsymbol{\varepsilon}}^T \mathbf{W} \hat{\boldsymbol{\varepsilon}}. \quad (9)$$

In nominal conditions, the pseudorange measurement error ε_i 's are modeled to be independent, Gaussian distributed random variables with zero mean and standard deviation δ_i , i.e., $\varepsilon_i \sim \mathcal{N}(0, \delta_i^2)$. In this condition, WSSE subjects to a chi-squared distribution with $n - 5$ degrees of freedom (DOF):

$$\text{WSSE} \sim \chi^2(n - 5). \quad (10)$$

The chi-squared test is applied to detect potential faults in the pseudorange measurements. Specifically, if WSSE exceeds a certain threshold T , a potential fault is claimed, and the positioning solution is assumed to be unreliable. On the contrary, if WSSE is below T , the measurements are declared fault-free, and the positioning solution is assumed to be valid. The threshold T is selected according to the predefined false alarm rate α , as shown below:

$$P(\text{WSSE} > T | H_0) = \alpha, \quad (11)$$

where H_0 is the fault-free hypothesis.

However, the Gaussian assumption about the pseudorange measurements may not always hold in real-world applications. Especially in urban scenarios, the GNSS signals can be affected by various environmental objects, such as high buildings and trees, which results in the reflection or diffusion of the GNSS signals, consequently distorting the measurement error distribution [10], [11]. These impacts have at least two severe outcomes.

- 1) The WLS solution relies heavily on the weighting matrix, which is directly related to the covariance of the measurement error distribution and is usually calibrated in the open sky condition [20], [30]. The mismatch of the weighting matrix and the measurement error distribution can substantially hinder the positioning accuracy in urban scenarios.
- 2) The chi-square test assumes that the measurement error in nominal conditions is Gaussian distributed. However, the measurement error distribution has been distorted by surroundings in the urban scenario, which cannot preserve the Gaussian properties [10]. In such cases, the test statistic WSSE could be substantially large so that the chi-squared test claims the existence of faults even if the pseudorange measurements are free of faults. Excluding these "faulty" measurements will reduce the measurement redundancy and cause the geometry to collapse [37], [38], further decreasing the positioning accuracy.

Therefore, a comprehensive approach combining adaptive error modeling and FDE is needed for GNSS positioning in urban scenarios, as illustrated in the following section.

III. SUBSPACE-BASED ADAPTIVE GMM ERROR MODELING AND FDE

The proposed method mainly consists of two parts, including the adaptive incremental pseudorange error modeling (AIPEM) algorithm and the adaptive fault detection and exclusion (AdaptiveFDE) algorithm. The fundamental idea is to adaptively maintain a set of GMM-based pseudorange error models identified by the elevation angle and carrier-to-noise ratio (C/N0) and apply them to exclude faulty measurements and solve positioning results. The main process is depicted in Figure 1. We first start with a short introduction to the Gaussian mixture model.

A. Gaussian Mixture Model

The GMM is a valuable statistical tool utilized in error modeling, especially for modeling heavy-tailed distributions that are commonly encountered in real-world scenarios [39]–[41]. It represents a probability distribution by combining multiple Gaussian components with different weights. The bimodal Gaussian mixture model (BGMM), a specific type of GMM, is particularly intriguing as it strikes an optimal balance between simplicity and flexibility, which offers the ability to accurately model both the central and extreme values of heavy-tailed distributions while minimizing the risks associated with overfitting and excessive parameters [40], [42]. A zero-mean BGMM can be formalized as

$$f(x) = p_1 \mathcal{N}(x; 0, \delta_1^2) + (1 - p_1) \mathcal{N}(x; 0, \delta_2^2), \quad (12)$$

where $\mathcal{N}(x; 0, \delta_1^2)$ and $\mathcal{N}(x; 0, \delta_2^2)$ are the PDF of the first and the second Gaussian component, δ_1 and δ_2 the corresponding standard deviations, and p_1 and $1 - p_1$ are the mixing weight of the two Gaussian components, respectively. In this work, we use the following expression to represent the distribution parameter of the zero-mean BGMM in (12):

$$\Theta = \{p_1, \delta_1, \delta_2\}. \quad (13)$$

The estimation of parameters in a GMM can be accomplished through Maximum Likelihood Estimation (MLE). A widely used method for obtaining MLE in GMM is utilizing the expectation–maximization (EM) algorithm [43]. This approach has garnered widespread popularity owing to its simplicity and effectiveness in accurately estimating the parameters of the GMM.

B. Adaptive Incremental Pseudorange Error Modeling

In urban conditions, goGPS is proposed to determine the uncertainty of GNSS pseudorange measurements based on the satellite geometry and the signal quality [31]. Specifically, the pseudorange measurement error is modeled by a Gaussian distribution, whose covariance is identified as a function of satellite elevation angle and carrier-to-noise ratio (C/N0). Given a measurement with the C/N0 and elevation angle

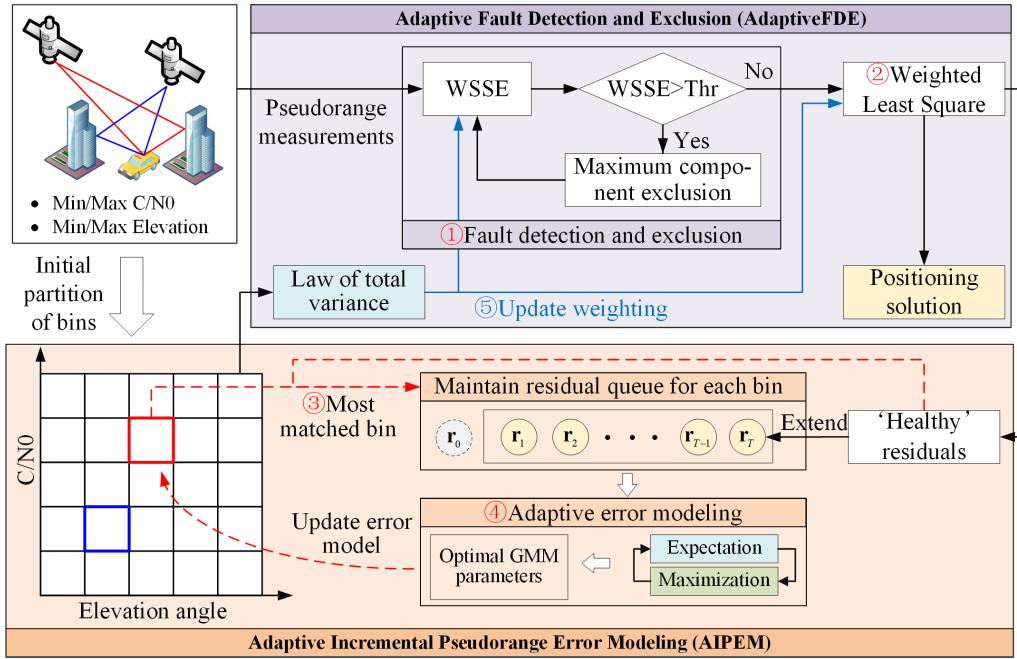


Fig. 1. Overview of the proposed subspace-based adaptive GMM error modeling and FDE method, which consists of AdaptiveFDE and AIPEM algorithms.

indicated as $CN0$ and e , respectively, its weight (the inverse of the covariance) can be calculated as follows [31]:

$$W(CN0, e) = \frac{1}{(\sin e)^2} \left(10^{-\frac{CN0 - s_1}{a}} \left(\left(\frac{A}{10^{-\frac{s_0 - s_1}{a}}} - 1 \right) \frac{CN0 - s_1}{s_0 - s_1} + 1 \right) \right), \quad (14)$$

where the s_1 parameter defines the C/N0 value for measurements that are considered “good”, the s_0 parameter defines the C/N0 value for which the function is forced to have the weight defined by the A parameter, and the a parameter defines the bending of the curve [31]. However, the error modeling process is static and primarily based on historical data, which may not capture the dynamic and diverse characteristics of urban environments.

Inspired by goGPS, we further proposed the adaptive incremental pseudorange error modeling (AIPEM) method to dynamically model the distribution of pseudorange error for real-time positioning systems in urban canyons, as shown in Algorithm 1. We assume that residuals obtained from the WLS solution with fault-free measurements (those that pass the chi-squared test) can represent the pseudorange measurement error [34]. The AIPEM algorithm can be described in the following steps:

(1) Initialization of bins based on subspace partition

Assume the step size of C/N0 is $\Delta\lambda$ and the step size of elevation angle is $\Delta\theta$, the GNSS pseudorange measurement space can be divided into $M \times N$ subspaces, each of which is represented with a bin. M and N represent the number of C/N0 bins and the number of elevation angle bins, respectively,

and can be calculated by (lines 3-4)

$$\begin{aligned} M &= \frac{\lambda_{\max} - \lambda_{\min}}{\Delta\lambda} \\ N &= \frac{\theta_{\max} - \theta_{\min}}{\Delta\theta}, \end{aligned} \quad (15)$$

where λ_{\max} and λ_{\min} are the maximum and minimum C/N0 specified by the dataset; θ_{\max} and θ_{\min} are the maximum and minimum elevation angle specified by the dataset.

For each bin, we initialize the measurement error distribution by implementing goGPS with default parameters [31], which produces a zero-mean Gaussian distribution (line 6). For example, the initialized error distribution parameter $\Theta_{i,j}^0$ in bin (i, j) is given by

$$\begin{aligned} (\delta_{i,j}^0)^2 &= \frac{1}{W(\lambda_{\min} + \frac{1}{2}i\Delta\lambda, \theta_{\min} + \frac{1}{2}j\Delta\theta)} \\ \Theta_{i,j}^0 &= \{0.5, \delta_{i,j}^0, \delta_{i,j}^0\}, \end{aligned} \quad (16)$$

where $W(\cdot)$ is defined in (14). In addition, a residual queue $R_{i,j}$ is initialized with $2T$ samples generated from the initialized error model (line 7), where T is the window size.

(2) Association of newly received residual with bins

Assume the measurement residuals at epoch k is

$$\hat{\varepsilon}^k = \{\hat{\varepsilon}_1^k, \hat{\varepsilon}_2^k, \dots, \hat{\varepsilon}_n^k\}. \quad (17)$$

For each residual component, we search the bin set to find the best-matched bin, which is defined as the bin where the elevation angle and C/N0 of the residual component fall into the subspace represented by it (lines 10-12). For example, the index of the best-matched bin of $\hat{\varepsilon}_1^k$ is given by

$$\begin{aligned} i &= \arg \min_i |\lambda_{\min} + i\Delta\lambda - \lambda_0| \quad \forall i \in [1, M] \\ j &= \arg \min_j |\theta_{\min} + j\Delta\theta - \theta_0| \quad \forall j \in [1, N], \end{aligned} \quad (18)$$

Algorithm 1 Adaptive Incremental Pseudorange Error Modeling (AIPEM)

Input:

Step size of C/N0: $\Delta\lambda$
 Step size of elevation angle: $\Delta\theta$
 Time window: T
 Measurement residuals at epoch k : $\hat{\varepsilon}^k = \{\hat{\varepsilon}_1^k, \hat{\varepsilon}_2^k, \dots, \hat{\varepsilon}_n^k\}$

Output:

Distribution parameters of pseudorange measurement error at each bin at epoch k : $\Theta_{i,j}^k$

```

1: Initialization (only once):
2:  $\lambda_{\max}, \lambda_{\min}, \theta_{\max}, \theta_{\min} \leftarrow$  Dataset
3:  $M \leftarrow \frac{\lambda_{\max} - \lambda_{\min}}{\Delta\lambda}$                                 ▷ Num. of C/N0 bins
4:  $N \leftarrow \frac{\theta_{\max} - \theta_{\min}}{\Delta\theta}$                                 ▷ Num. of Ele bins
5:  $\Xi = \{(i, j) | \forall i \in [1, M], j \in [1, N]\}$     ▷ Indices of bins
6:  $(\delta_{i,j}^0)^2 = \frac{1}{W(\lambda_{\min} + \frac{1}{2}i\Delta\lambda, \theta_{\min} + \frac{1}{2}j\Delta\theta)}$ , ▷ goGPS weighting
    $\Theta_{i,j}^0 = \{0.5, \delta_{i,j}^0, \delta_{i,j}^0\} \quad \forall (i, j) \in \Xi$ 
7:  $R_{i,j} \leftarrow 2T$  samples generated from  $\Theta_{i,j}^0 \quad \forall (i, j) \in \Xi$           ▷ Residual queue
8:  $\eta_{i,j} \leftarrow 0 \quad \forall (i, j) \in \Xi$                                      ▷ Update flag
9: for  $s \leftarrow 1$  to  $n$  do
10:    $\lambda_0, \theta_0 \leftarrow$  The GNSS signal associated with  $\hat{\varepsilon}_s^k$ 
11:    $i = \arg \min_i |\lambda_{\min} + i\Delta\lambda - \lambda_0| \quad \forall i \in [1, M]$ 
12:    $j = \arg \min_j |\theta_{\min} + j\Delta\theta - \theta_0| \quad \forall j \in [1, N]$ 
13:    $R_{i,j}.push(\hat{\varepsilon}_s^k)$ 
14:    $\eta_{i,j} = 1$                                          ▷ Bin  $(i, j)$  is updated
15: end for
16: for each  $(i, j) \in \Xi$  do
17:   if  $\eta_{i,j} \neq 0$  then
18:      $\Theta_{i,j}^k \leftarrow EM(R_{i,j})$                       ▷ Fit GMM with EM
19:   else
20:      $R_{i,j}.push(2$  samples generated from  $\Theta_{i,j}^{k-1})$           ▷ Realizaton of incremental update
21:      $\Theta_{i,j}^k \leftarrow \Theta_{i,j}^{k-1}$ 
22:   end if
23:    $R_{i,j}.pop(\text{out-of-date samples})$ 
24: end for

```

where λ_0 and θ_0 is the C/N0 and elevation angle of the pseudorange measurement associated with $\hat{\varepsilon}_1^k$.

(3) Maintenance of residual queue and error model

In each bin, we maintain the residual queue $R_{i,j}$ by feeding the matched residual (line 13) and then apply the EM algorithm [43] to fit a zero-mean BGMM distribution for data in the updated residual queue of the bin (line 18). If a bin does not receive new residual data, we use the latest distribution parameter to generate a fixed number of samples to expand its residual queue (line 20), while the error model remains unchanged (line 21). This operation ensures the incremental update of the pseudorange error model. Specifically, the majority of the residual data inherits from the historical residual distribution, and newly acquired residual data will only have a limited impact on the overall distribution. As a result, the update of the residual distribution is guaranteed to avoid sudden changes or mutations. Moreover, we drop

out-of-date elements in each bin so that the time span of the queue does not exceed a predefined window size T (line 23). Specifically, we remove those residuals whose timestamp is less than $k - T$, where k is the current epoch.

The AIPEM method can incrementally update the pseudorange error distribution with adaptation to the environment, providing environment-aware information for positioning and fault detection, which is illustrated in the following section.

C. Adaptive Fault Detection and Exclusion for Real Time Differential Positioning

Assume the zero-mean BGMM distribution parameter in bin (i, j) at epoch k is

$$\Theta_{i,j}^k = \{p_1^k, \delta_1^k, \delta_2^k\}, \quad (19)$$

According to the law of total covariance [44], the total variance of the BGMM $\Theta_{i,j}^k$ is given by

$$(\delta_*^k)^2 = p_1^k(\delta_1^k)^2 + (1 - p_1^k)(\delta_2^k)^2 \quad (20)$$

The total variance reflects the overall uncertainty of the pseudorange error modeled by GMM, which can be utilized to adjust the weighting matrix in the WLS solution scheme and the chi-squared test process. We implement this idea in the real-time GNSS positioning system in urban canyons and propose the adaptive fault detection and exclusion (AdaptiveFDE) algorithm, as shown in Algorithm 2. The AdaptiveFDE algorithm can be described in the following steps:

(1) Construction of weighting matrix

Assume the s th pseudorange measurement at epoch k is ρ_s^k , and its elevation angle and C/N0 are θ_s and λ_s , respectively, then the associated bin (i, j) of ρ_s^k is determined by (lines 4-5)

$$i = \arg \min_i |\lambda_{\min} + i\Delta\lambda - \lambda_s| \quad (21a)$$

$$j = \arg \min_j |\theta_{\min} + j\Delta\theta - \theta_s|. \quad (21b)$$

By using the distribution parameter $\Theta_{i,j}^{k-1}$ in bin (i, j) at epoch k , we can calculate the error covariance $(\delta_s^k)^2$ of the measurement ρ_s^k by employing the law of total covariance (line 6), as shown in (20). Then the inverse of $(\delta_s^k)^2$ is taken as the (s, s) element of the weighting matrix (line 7). By performing the above operation for all pseudorange measurements, we can obtain the weighting matrix \mathbf{W} at epoch k .

(2) Fault detection and exclusion

The weighting matrix \mathbf{W} is utilized to calculate the residual vector $\hat{\varepsilon}^k$ with (7) and the test statistic $WSSE_k$ with (9) at epoch k (lines 9-10). A chi-squared test is then conducted to examine whether $WSSE_k$ exceeds the predefined threshold $\chi^{2-1}(1 - \alpha)$, where α is the significant level of the chi-squared test and $\chi^2(\cdot)$ is the chi-squared distribution.

If the test statistic exceeds the threshold, a sequential measurement exclusion process is performed. Specifically, we first calculate the normalized residual vector as follows:

$$\mathbf{W}^{\frac{1}{2}}\hat{\varepsilon}^k. \quad (22)$$

The index of the absolute maximum element in $\mathbf{W}^{\frac{1}{2}}\hat{\varepsilon}^k$ can be easily found and denoted as s (line 12). Then the s th

Algorithm 2 Adaptive Fault Detection and Exclusion (AdaptiveFDE)

Input:

Step size of C/N0: $\Delta\lambda$
 Step size of elevation angle: $\Delta\theta$
 Distribution parameters at epoch $k - 1$: $\Theta_{i,j}^{k-1} \forall (i,j) \in \{(i,j) | \forall i \in [1, M], j \in [1, N]\}$
 Measurements at epoch k : $\rho^k = \{\rho_1^k, \rho_2^k, \dots, \rho_n^k\}$
 False alarm rate: α

Output:

Positioning solution at epoch k : \hat{x}^k
 Finalized test statistic at epoch k : $WSSE_k$

- 1: Initialization: $\mathbf{W} = \mathbf{I}_{n,n}$
- 2: **for** $s \leftarrow 1$ to n **do**
- 3: $\lambda_s, \theta_s \leftarrow$ The GNSS signal associated with ρ_s^k
- 4: $i = \arg \min_i |\lambda_{\min} + i\Delta\lambda - \lambda_s| \forall i \in [1, M]$
- 5: $j = \arg \min_j |\theta_{\min} + j\Delta\theta - \theta_s| \forall j \in [1, N]$
- 6: $(\delta_s)^2 \leftarrow$ Equation (20) with $\Theta_{i,j}^{k-1}$
- 7: $w_{s,s} = \frac{1}{(\delta_s)^2}$ \triangleright Set (s, s) element of \mathbf{W}
- 8: **end for**
- 9: $\hat{\varepsilon}^k \leftarrow$ Equation (7) with ρ^k and \mathbf{W}
- 10: $WSSE_k = \hat{\varepsilon}^{kT} \mathbf{W} \hat{\varepsilon}^k$
- 11: **while** $WSSE_k > \chi^{2-1}(1 - \alpha)$ and $|\hat{\varepsilon}^k| > 5$ **do**
- 12: $s \leftarrow$ Index of absolute maximum element in $\mathbf{W}^{\frac{1}{2}} \hat{\varepsilon}^k$
- 13: Remove s th element of ρ^k
- 14: Remove s th row and s th column of \mathbf{W}
- 15: $\hat{\varepsilon}^k \leftarrow$ Equation (1) to (7) with ρ^k and \mathbf{W}
- 16: $WSSE_k = \hat{\varepsilon}^{kT} \mathbf{W} \hat{\varepsilon}^k$
- 17: **end while**
- 18: $\hat{x}^k \leftarrow$ Equation (1) to (6) with ρ^k and \mathbf{W}
- 19: Execute Algorithm 1 to obtain $\Theta_{i,j}^k$ with $\hat{\varepsilon}^k$

element in ρ^k is eliminated (line 13) and the s th row and s th column of \mathbf{W} are also eliminated (line 14). Subsequently, the residual vector and the test statistic are re-calculated (lines 15-16). Such the process is repeated until either 1) the test statistic is no larger than the threshold or 2) the minimum sufficient measurements requirement is no longer satisfied. In our positioning system, the minimum number of measurements for implementing WLS is 5 (three unknown variables associated with position and two unknown variables associated with the receiver clock bias from two satellite constellations, i.e., GPS and BeiDou).

(3) Solving position and updating error distribution

The remaining measurements are utilized to solve the GNSS receiver's position by WLS (line 18), while the associated residuals are fed to Algorithm 1 to calculate the distribution parameter at epoch k (line 19).

IV. EXPERIMENTAL RESULTS

A. Experiment Setup

To examine the performance of the proposed method, we conducted experiments on two datasets, including the KLT dataset and the TST dataset. The KLT dataset is collected in a residual area with low-rise houses, as shown in Fig. 2(a).

TABLE I
 THE SETTING OF PARAMETERS ON EACH DATASET

Parameters	Value
GNSS constellations	GPS, Beidou
C/N0 step $\Delta\lambda$	2 dB Hz
Elevation angle step $\Delta\theta$	5°
Time window T	60 s
False alarm rate α	0.05

A surveying vehicle is employed to drive along the designed track for data collection. Fig. 2(c) shows the setup of the sensors on the surveying vehicle, where the U-blox Zed F9P is employed to collect L1 GPS and BEIDOU signals at a frequency of 1 Hz. The ground truth of the receiver position is provided by a NovAtel SPAN-CPT, a GNSS RTK/inertial navigation system (INS) integrated navigation system. The TST dataset is fetched from UrbanNav [45], an open-sourcing localization dataset. The TST dataset is collected by the EVK-M8T in an urbanized area in Hong Kong with a limited sky-view due to the existence of high buildings, as shown in Fig. 2(b). Similar to the KLT dataset, the ground truth of the receiver position is provided by NovAtel SPAN-CPT (1 Hz).

Fig. 3(a) and (b) plot the distribution of the number of GNSS signals at each epoch on the two datasets, respectively. Overall, the number of GNSS signals at each epoch of the KLT dataset is significantly larger than that of the TST dataset, with a mean value of 15.88. The primary reason is that the TST dataset is collected in a medium urban scenario where high buildings are likely to block some GNSS signals. This phenomenon is further validated by drawing the quantile-quantile (QQ) plot of the GNSS pseudorange measurement error, as shown in Fig. 3(e) and (f). The QQ plot shows the quantile of error distribution with the equivalent standard normal quantile, such that Gaussian distributed error data exhibit a straight line. In these QQ plots, we choose the pseudorange measurement error with elevation angles observed from 30° to 40°. As can be seen, the measurement error from both datasets shows a heavy-tailed characteristic, whereas the measurement error from the TST dataset displays a more pronounced phenomenon. Being significantly affected by the surroundings, GNSS signals in the TST dataset can be refracted and reflected; thereby, their distribution is distorted and demonstrates the long tail phenomenon. Therefore, it is reasonable to adopt GMM to characterize error profiles of pseudorange measurements on these two urban datasets. Fig. 3(c) and (d) show the absolute pseudorange measurement errors against the elevation angle and C/N0 related to the receiver on the two datasets, respectively, which helps us to identify the range of elevation angles and C/N0. The setting of the step size, time window, and false alarm rate are also listed in Table I.

B. Performance of Adaptive Error Modeling

On both datasets, we implement the AdaptiveFDE algorithm. We first evaluate the effectiveness of the proposed

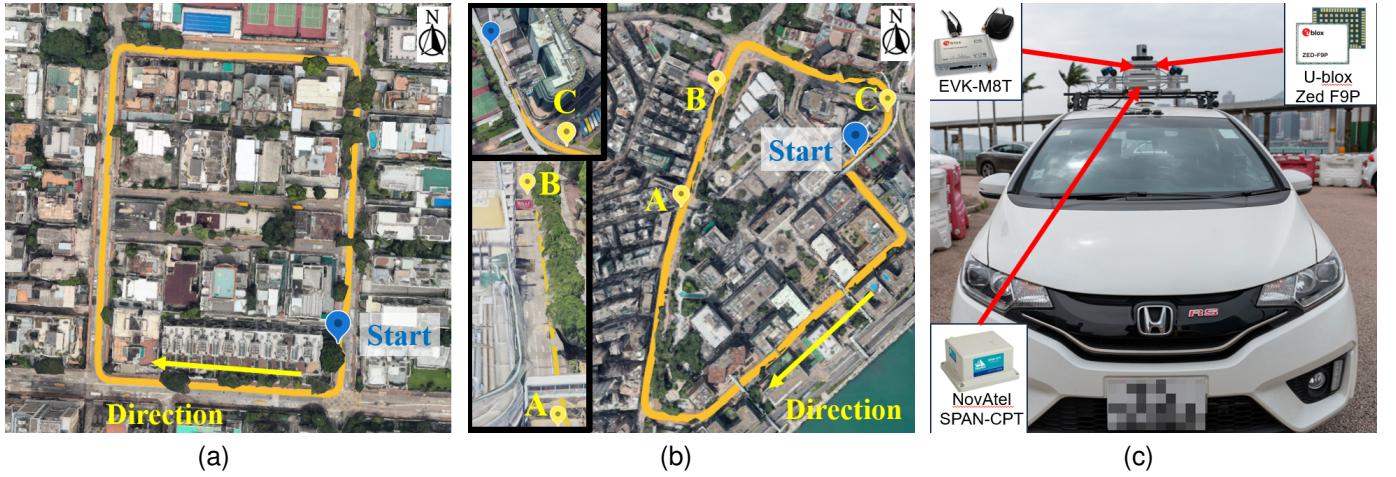


Fig. 2. The trajectory of the surveying vehicle on (a) the KLT dataset and (b) the TST dataset. (c) The installation of the GNSS receiver on the surveying vehicle.

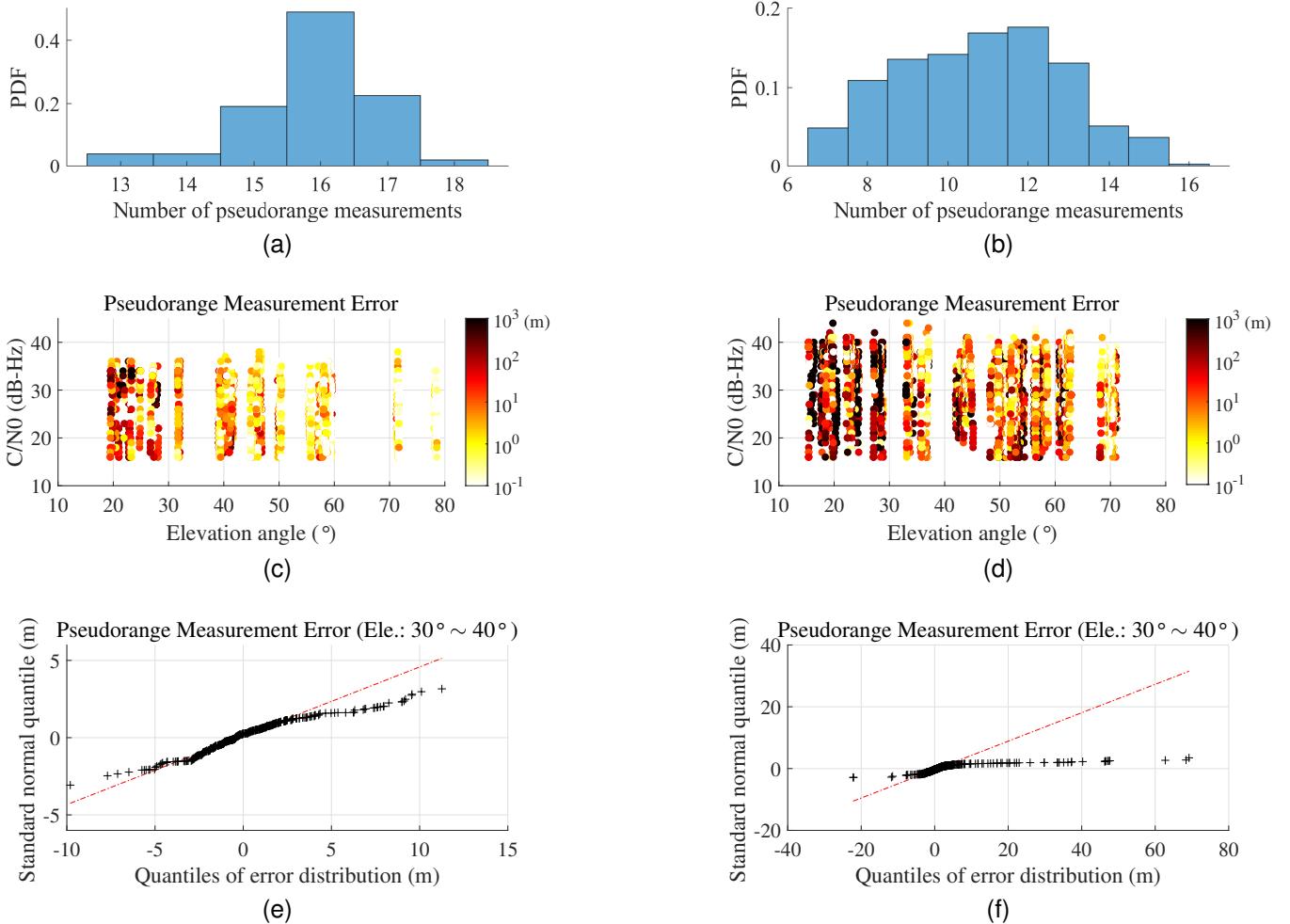


Fig. 3. The distribution of the number of GNSS signals at each epoch on (a) the KLT dataset and (b) the TST dataset. The absolute pseudorange measurement errors against the elevation angle and C/N0 related to the receiver on (c) the KLT dataset and (d) the TST dataset. The color bar represents the magnitude of the absolute error and is plotted in the logarithmic scale. The QQ plot of GNSS pseudorange measurement error for elevation angles observed from 30° to 40° on (e) the KLT dataset and (f) the TST dataset.

method in the error modeling task. Regarding the KLT dataset, Fig. 4(a) plots the standard deviation difference between two

Gaussian components against time in each bin. Specifically, we assume that the 2nd Gaussian component has a larger standard

deviation than the 1st component without loss of generality. Therefore, the standard deviation difference in Fig. 4(a) is defined as the standard deviation of the 2nd Gaussian component subtracted from that of the 1st component, consequently taking a non-negative value. In addition, each row in Fig. 4(a) corresponds to a bin. The row number k of $\text{Bin}(i, j)$ is

$$k = i * \frac{\lambda_{\max} - \lambda_{\min}}{\Delta\lambda} + j. \quad (23)$$

As can be seen, the standard deviation difference in each bin changes with time, demonstrating the dynamic nature of our error modeling process. Nevertheless, we observe that the value in two regions, marked with dashed rectangular, almost approaches zero, suggesting that errors in these bins have Gaussian distributions. Since the KLT dataset is collected in an almost open-sky condition, it is not surprising that some pseudorange measurement errors can maintain normality.

However, such a phenomenon is rare in the TST dataset, as shown in Fig. 4(b). The standard deviation difference in most bins shows a consistently time-varying property. To examine the adaptability of the proposed algorithm to changing environments, we choose a specific bin that exhibits significant changing modes for further study, i.e., $\text{Bin}(\lambda = 28 \sim 30 \text{ dB-Hz}, \theta = 40 \sim 45^\circ)$. Fig. 4(c) and Fig. 4(d) plot the distribution parameters of this selected bin against time. Interestingly, the curve is naturally divided into four parts by three epochs: t_A , t_B , and t_C . The locations corresponding to the three epochs are marked in Fig. 2(b). On the left side of path AB, we observe the presence of tall buildings, while on the right side is a row of trees. This configuration differs from the situation prior to point A, where there is no tree near the left of the path. Assuming that the vehicle on the path AB is $d = 2 \text{ m}$ far from the trees, we can calculate the minimum height of trees, h_{\min} , that can block GNSS signals with the elevation angle of $40 \sim 45^\circ$ by

$$h_{\min} = x \tan(40 * \pi / 180) = 1.68 \text{ m}. \quad (24)$$

By and large, the height of trees along the left side of path AB exceeds h_{\min} ; thereby, these trees can substantially affect the GNSS signals with the elevation angle of $40 \sim 45^\circ$. The changes of distribution parameters within $t_A \sim t_B$ coincide with the qualitative analysis, where the standard deviation of the 2nd component and the associated mixing weights ($1 - p_1$) increase during this period, which indicates capturing a heavy-tailed error distribution. The impacts of trees and high buildings are temporally diminished along the path BC, which is a wide street. Referring to Fig. 4(c) and Fig. 4(d), we observe a notable rise in p_1 during this period, while the standard deviations of both components do not exhibit significant variations. This phenomenon indicates a reduction in measurement variance, implying that the impact of the surrounding environment on GNSS signals has gradually diminished. Upon departing from point C, the vehicle proceeds onto a narrow street beneath a pedestrian footbridge, as shown in Fig. 2(b). This footbridge, combined with high buildings lining the road, significantly affects the GNSS signal, leading to a considerable increase in measurement uncertainty. Our adaptive error modeling accurately captures this phenomenon,

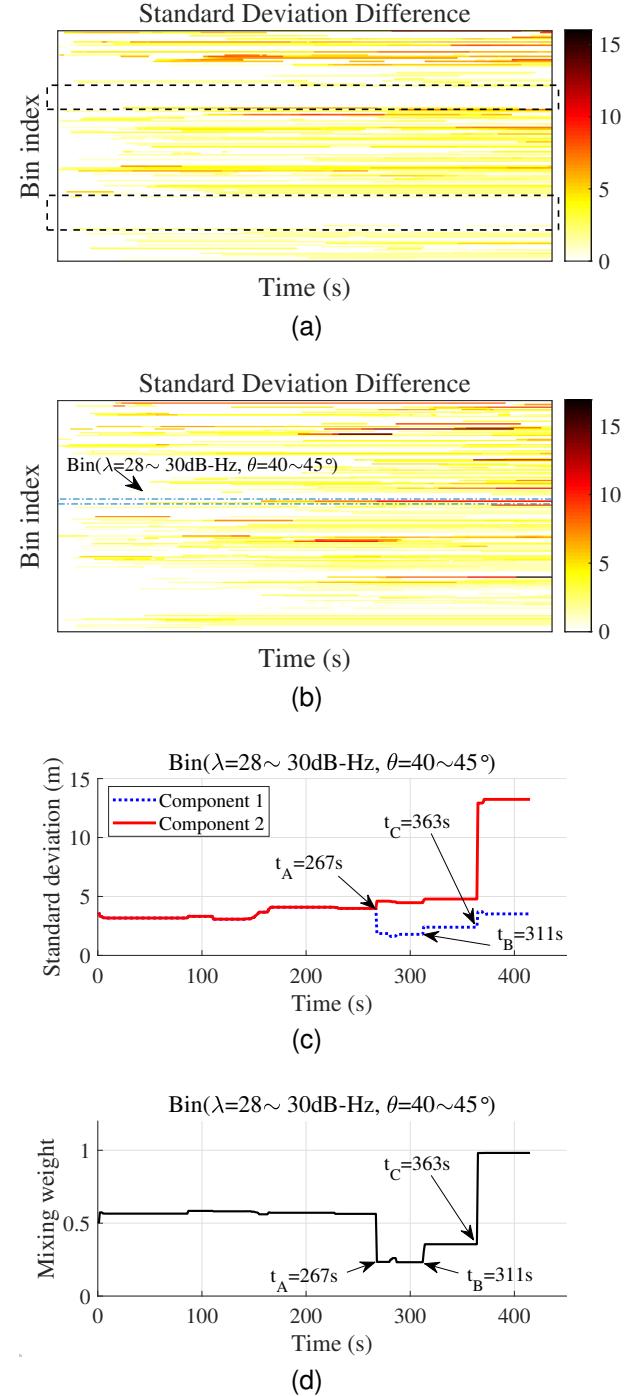


Fig. 4. The difference of standard deviation between two Gaussian components, i.e., $\delta_2 - \delta_1$, against time in each bin on (a) the KLT dataset and (b) the TST dataset. Each row represents a bin. The evolution of (c) standard deviation and (d) mixture weight parameters of the GMM distribution in the $\text{Bin}(\lambda = 28 \sim 30 \text{ dB Hz}, \theta = 40 \sim 45^\circ)$ on the TST dataset.

as evidenced by a substantial rise in the standard deviations of both components during this period. Notably, the standard deviation of the 2nd component surpasses that of the 1st component, indicating a pronounced heavy-tailed distribution of the measurement error. The statistics of the computation time of the error modeling process in both datasets are given in Appendix B.

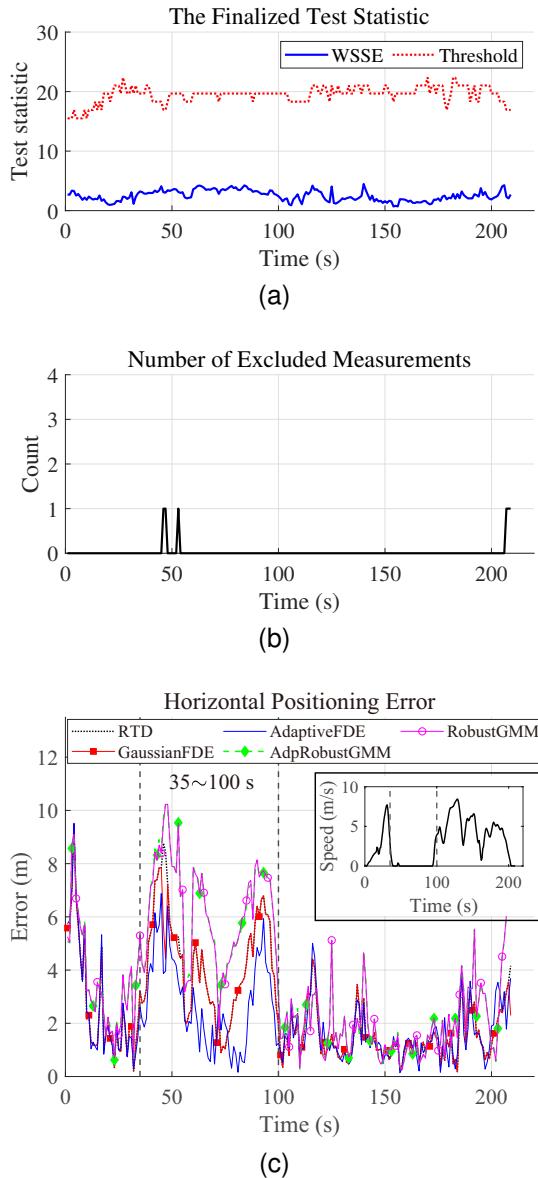


Fig. 5. (a) The finalized test statistic of AdaptiveFDE after excluding all potential faults on the KLT dataset. (b) The number of excluded measurements at each epoch on the KLT dataset. (c) The horizontal positioning error of RTD, GaussianFDE, RobustGMM, AdpRobustGMM, and AdaptiveFDE in the ENU coordinates on the KLT dataset.

C. Performance of Fault Detection and Positioning

The AdaptiveFDE algorithm utilizes the adaptive error model to construct the test statistic and conduct the sequential exclusion of faulty measurements. The finalized test statistic after fault exclusion on the KLT dataset is plotted against time in Fig. 5(a). Since the KLT dataset has relatively healthy measurements, the test statistic remains considerably lower than the threshold, with faults declaring only at a few specific timestamps, as shown in Fig. 5(b). In this condition, the adaptive error modeling mainly benefits the WLS solution by consistently adjusting weights.

We further compare the positioning accuracy of the proposed AdaptiveFDE method with RTD (Section II-A), GaussianFDE (Section II-B) [20], robust estimation with GMM

error modeling (RobustGMM) [35], and robust estimation with adaptive GMM error modeling (AdpRobustGMM) [34]. To implement RobustGMM, we first calculate and aggregate the measurement errors in the KLT dataset based on receiver positions. Then, these aggregated measurement errors are utilized to fit a two-component GMM, which is regarded as the nominal error model for robust estimation. Note that all GNSS measurements share the same nominal error model in RobustGMM, and this error model would not change with time. For the AdpRobustGMM, we implement the algorithm in [34] with the default parameters, where the window size of residuals is 60 and the number of Gaussian components in the GMM is 2. Although the error model is adaptively tuned in AdpRobustGMM with time, all GNSS measurements at each time epoch share the same error model.

As shown in Fig. 5(c), the positioning error of AdaptiveFDE is smaller than the other four methods nearly at all time epochs. RobustGMM and AdpRobustGMM yield the largest positioning error, which is even larger than that of RTD. This is because AdpRobustGMM and RobustGMM assume that the GNSS measurements obtained from different satellite-user geometry have the same error distribution, which is impractical in real-world applications. The GaussianFDE exhibits a performance that closely resembles that of the AdaptiveFDE. However, the positioning error of GaussianFDE is significantly larger than that of AdaptiveFDE during the period of $35 \sim 100$ s. The thumbnail plot in Fig. 5(c) shows that the vehicle is almost static during this period, potentially explaining why AdaptiveFDE has a better positioning performance than GaussianFDE. During this period, the GNSS signals received by the vehicle's receiver have a fixed range of elevation angles since the vehicle's surroundings, such as buildings and trees, are unchanging. Therefore, a specific set of bins in the AIPEM algorithm can consistently obtain residual data, gradually adjusting their error model to describe the environmental effects on the pseudorange measurements. As a result, the AdaptiveFDE algorithm can adopt more suitable weights than the GaussianFDE algorithm, demonstrating significantly lower positioning error during this period.

It is also important to analyze the performance of AdaptiveFDE outside this period. Before 35 s, the bins in the AIPEM algorithm have not collected enough residual data to describe the heavy-tailed properties of the measurement errors; therefore, the AdaptiveFDE takes almost the same weights as the GaussianFDE during this period, showing no difference in the positioning errors. The performance of AdaptiveFDE is almost the same as GaussianFDE after 100 s, which is similar to the case before 35 s; However, the underlying rationale behind this similarity differs. After 100 s, the vehicle travels along a street in a low-rise residential area, where the heights of the houses are fairly uniform. During this movement, each bin is unlikely to acquire sufficient data to characterize the subtle variations in the impact of the surroundings on GNSS signals. Therefore, the error model in most bins is unaltered and retains the distribution parameter initialized by goGPS. As a result, the positioning error of AdaptiveFDE closely approximates that of GaussianFDE. Table II lists the mean positioning error of RTD, GaussianFDE, RobustGMM,

AdpRobustGMM and AdaptiveFDE. As can be seen, the value of AdaptiveFDE is only 2.14 m, which is around 16 % smaller than that of GaussianFDE.

TABLE II
THE MEAN ABSOLUTE HORIZONTAL POSITIONING ERROR

Methods	KLT Dataset	TST Dataset
RTD	2.60 m	12.58 m
GaussianFDE	2.55 m	9.67 m
RobustGMM	3.62 m	17.43 m
AdpRobustGMM	3.62 m	17.44 m
AdaptiveFDE	2.14 m	8.79 m

Regarding the TST dataset, we plot its finalized test statistic after fault exclusion in Fig. 6(a), which shows a different mode to that of the KLT dataset. Since the TST dataset is collected in the medium urban scenario, the GNSS pseudorange measurement can be severely affected by the surroundings and consists of considerable faults. In this case, the AdaptiveFDE algorithm demonstrates its ability to detect and exclude potential faults, as suggested by Fig. 6(b). When all faults are excluded, the finalized test statistic is moderately smaller than the threshold. Fig. 6(c) plots the trajectory of the positioning results of RTD, GaussianFDE, RobustGMM, AdpRobustGMM, and AdaptiveFDE, while the ground truth (GT) is also plotted in this figure. Compared to RTD, RobustGMM, and AdpRobustGMM, which do not apply fault detection and exclusion, AdaptiveFDE produces a smoother positioning trajectory with smaller positioning errors. However, GaussianFDE also demonstrates comparative performance in positioning. To compare GaussianFDE and AdaptiveFDE, we calculate the difference of horizontal positioning error between AdaptiveFDE and GaussianFDE in the east-north-up (ENU) coordinates and plot it against time in Fig. 6(d), where the negative value indicates that AdaptiveFDE has a smaller positioning error than GaussianFDE. As can be seen, AdaptiveFDE exhibits smaller positioning errors than GaussianFDE most of the time. A quantitative comparison is given in Table II that the mean absolute horizontal positioning error of AdaptiveFDE is only 8.79 m, while this figure of GaussianFDE, RobustGMM, and AdpRobustGMM is 9.66 m, 17.43 m, and 17.44 m, respectively. By incorporating adaptive error modeling in both the weighted least squares (WLS) and fault detection and exclusion processes, AdaptiveFDE achieves a notable 9 % reduction in positioning error compared to GaussianFDE. Indeed, the improvement in positioning performance can partially be attributed to the effects of FDE on measurement error distribution. Due to the existence of unmodelled errors in measurements, the distribution of measurement error tends to be biased and has heavy tails. By applying fault detection and exclusion, we can improve the consistency between measurements and refine the bias and heavy tail issues in the error distribution. The effects of FDE on measurement error distribution are discussed in Appendix A.

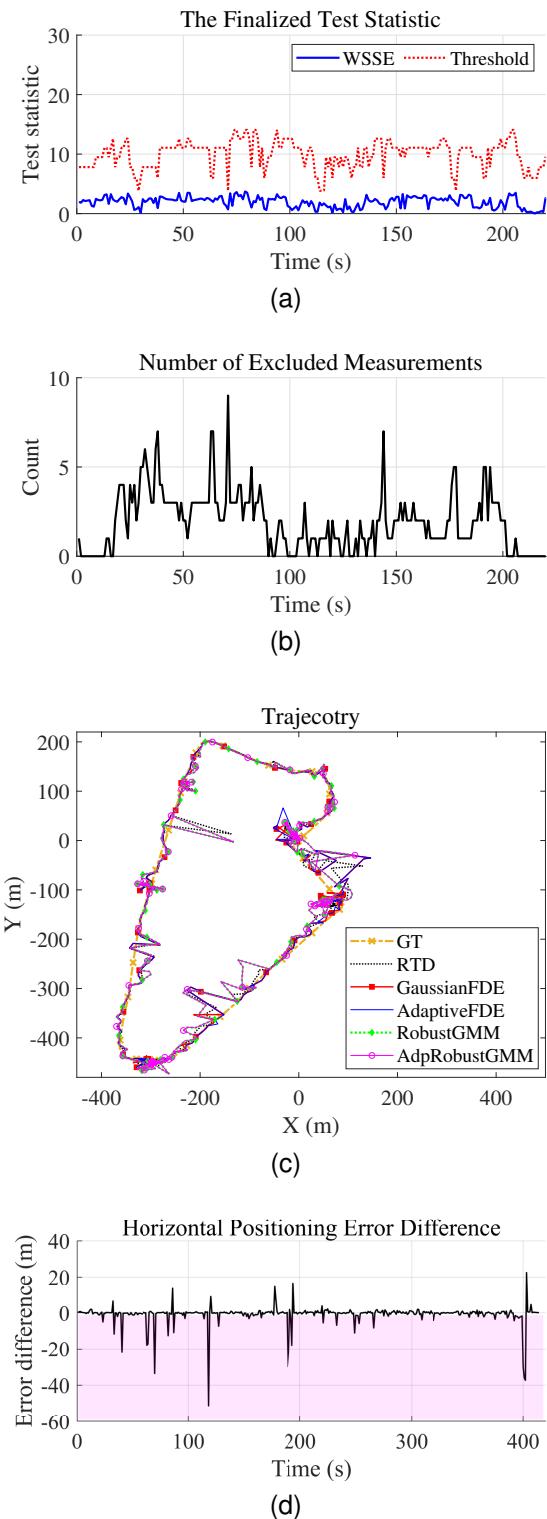


Fig. 6. (a) The finalized test statistic of AdaptiveFDE after excluding all potential faults on the TST dataset. (b) The number of excluded measurements at each epoch on the TST dataset. (c) The trajectory of the positioning results of RTD, GaussianFDE, RobustGMM, AdpRobustGMM, and AdaptiveFDE on the TST dataset. (d) The difference in horizontal positioning error between AdaptiveFDE and GaussianFDE in the ENU coordinates on the TST dataset. The shaded area indicates that AdaptiveFDE has a smaller positioning error than GaussianFDE.

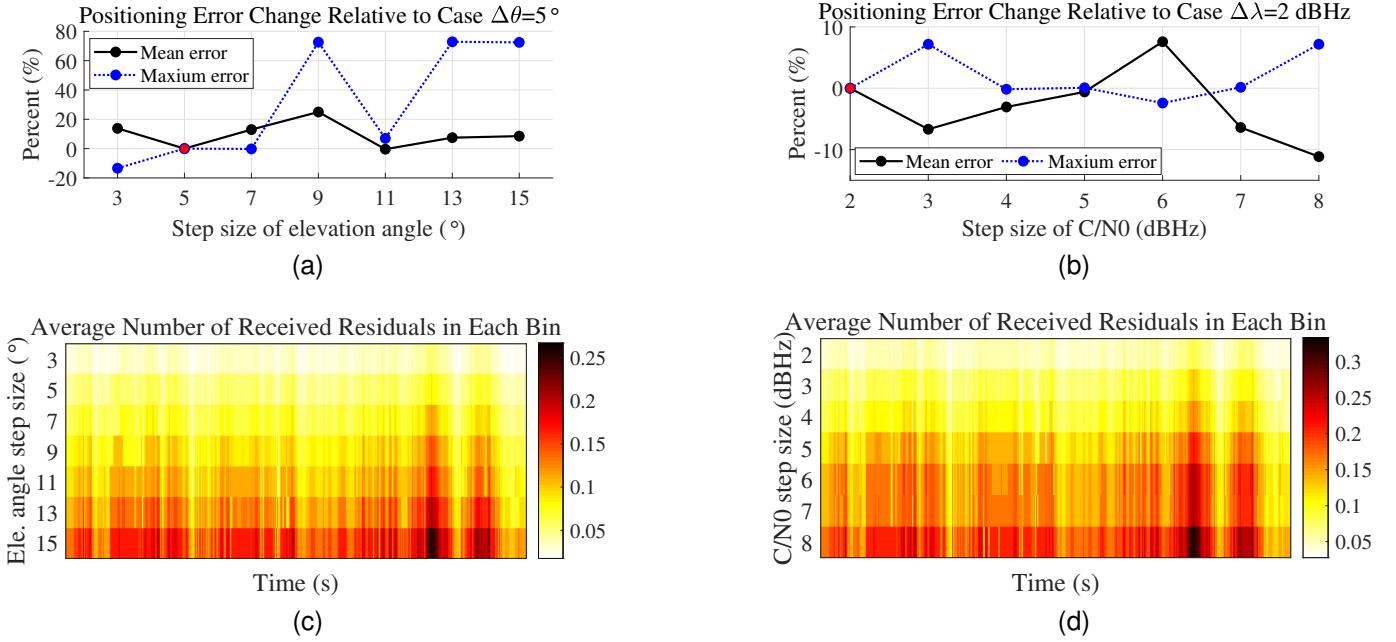


Fig. 7. The change of mean and maximum horizontal positioning error relative to the benchmark against the step size of (a) elevation angle and (b) C/N0. The benchmark is marked with a red dot. The evolution of the average number of received residuals in each bin against the step size of (c) elevation angle and (d) C/N0. All experiments are conducted in the TST dataset.

V. DISCUSSION

A. Impacts of Step Size

In the proposed AdaptiveFDE and AIPEM algorithm, users need to specify the step size of elevation angle $\Delta\theta$ and the step size of C/N0 $\Delta\lambda$ for partitioning data bins. To examine their effects on positioning performance, we conducted controlled experiments on the TST dataset. Specifically, we set the value of $\Delta\theta$ to $3^\circ, 7^\circ, 9^\circ, 11^\circ, 13^\circ$, and 15° while keeping all other parameters consistent with Table I. We calculate the change in mean and maximum positioning error for each setting relative to the benchmark ($\Delta\theta = 5^\circ$) and plot the results in Fig. 7(a), where a positive value indicates the increase in positioning error compared to the benchmark (marked with a red dot). Obviously, the maximum positioning error increases with the step size of elevation angle, implying that the error distribution is notably affected by the elevation angle. If we employ a larger elevation angle step size for the bin's partition, the collected residual data in each bin would exhibit greater heterogeneity, which hinders the effectiveness of error modeling.

The increase in the step size of elevation angle also has a negative impact on the overall positioning performance. However, this negative effect is partially offset by the additional benefits, i.e., increased number of received residuals per unit time, that come with increasing the step size. Fig. 7(c) plots the evolution of the average number of received residuals in each bin against the step size of elevation angle. With the increase in step size, the average number of received residuals in each bin per unit time significantly increases. Each bin will have more residual samples for error modeling, which is beneficial for capturing the common pattern inside the data. Therefore, it is observed that the mean positioning error does not experience

a sudden increase but instead fluctuates within the range of 0 to 20 % when the step size of the elevation angle increases, as shown in Fig. 7(a). Nevertheless, we recommend using a relatively small step size of elevation angle. In this case, 5° is the optimal choice, which balances the positioning error and the data availability within each bin.

Regarding the impacts of C/N0 step size, we compare the settings of 3 dB Hz, 4 dB Hz, 5 dB Hz, 6 dB Hz, 7 dB Hz, and 8 dB Hz with the benchmark (2 dB Hz). As shown in 7(b), both the mean and maximum positioning error fluctuate within the range of -10 % to 10 % with the increase of C/N0 step size from 2 dB Hz to 6 dB Hz. This phenomenon is quite anti-intuitive since C/N0 is always thought to be an indicator to justify the quality of GNSS signals. However, considering that the TST dataset is collected in the medium urban scenario, the impacts of surroundings on C/N0 are quite complex. Apart from the reflection and obstruction by buildings, atmospheric conditions and radio frequency interference (RFI) can also cause substantial impacts on C/N0. It is likely that a nonlinear relationship exists between C/N0 and the patterns in GNSS pseudorange measurement errors. Since adjusting the C/N0 step size is a linear operation, it may have limited effect in separating various patterns within the error data. However, when the C/N0 step size is larger than 6 dB Hz, a downward trend in the mean positioning error becomes apparent as the step size increases. Enlarging the C/N0 step size increases the average number of received residuals within each bin per unit time, as shown in Fig. 7(d). As the level of elevation angle specified by each bin remains constant regardless of the change in C/N0 step size, the augmented samples obtained by applying large C/N0 step size help identify the common patterns within the error data regarding the elevation angle.

Consequently, the mean positioning error decreases. Interestingly, the maximum positioning error increases during this interval, indicating that the influence of C/N0 step size is not negligible. Further investigation is required to understand the impacts of C/N0 step size. In this case, we recommend using 4 dB Hz as C/N0 step size to balance the mean and maximum positioning error.

B. Impacts of Time Window

In the AIPEM algorithm, time window T determines the length of the residual queue, which determines the memory size of history information. Fig. 8 plots the impacts of time window on the positioning performance. Similar to the experiments in Section V-A, we keep all parameters, except time window, the same as the benchmark in Table I, and implement the AdaptiveFDE on the TST dataset with different settings of time window. As suggested by Fig. 8, the maximum positioning error exhibits an upward trend as the window size increases, implying that a long time window can impede the sensitivity of the error modeling algorithm to environment changes. Nevertheless, the mean positioning error is relatively stable across all settings of time window. This can be explained by the “average effects” of a long time window. Although a longer time window may not respond immediately to sudden environmental changes, its ability to maintain a large amount of historical data allows for a smoother update of the error model. This results in a more generalized error model, leading to competitive performance regarding mean positioning error.

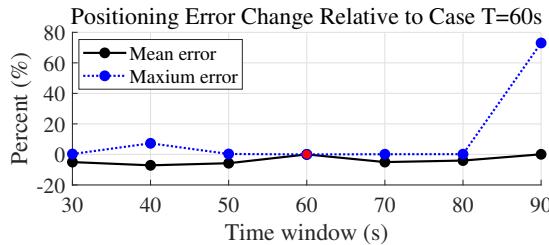


Fig. 8. The change of mean and maximum horizontal positioning error relative to the benchmark (the red dot) against the time window on the TST dataset.

VI. CONCLUSION

In this work, we propose the subspace-based adaptive GMM error modeling and FDE method for GNSS positioning in urban canyons. This method adaptively maintains a set of pseudorange error models identified by the elevation angle and carrier-to-noise ratio (C/N0), which are employed to construct the fault detection and exclusion process and determine the positioning solution. Specifically, we propose the adaptive incremental pseudorange error modeling (AIPEM) algorithm to dynamically model the distribution of pseudorange error for real-time positioning systems in urban scenarios. Based on the adaptive error model, we further propose the adaptive fault detection and exclusion (AdaptiveFDE) algorithm for RTD positioning to dynamically adjust the weighting matrix in the FDE process and the WLS solution scheme.

The performance of the proposed method is evaluated on two urban datasets, including the KLT dataset representing slightly urbanized environments and the TST dataset representing a medium urbanized scenario. Results show that the evolution of error models has a strong relationship with surroundings, suggesting that the proposed method has the ability to sense environmental changes. In addition, compared to the conventional FDE method with Gaussian assumptions, the mean positioning error is reduced by 16 % and 9 % in the KLT dataset and the TST dataset, respectively, showing the power of adaptive error modeling in enhancing positioning accuracy. We further discuss the impact of step size (elevation angle and C/N0) and time window parameters of two algorithms on positioning with controlled experiments, providing practical guidelines for GNSS positioning in urban environments.

This study has several limitations, which also point out future research directions. In constructing the AIPEM algorithm, we limit the error model to the zero-mean GMM distribution for simplicity. However, the multipath and NLOS effect caused by urban environments could have complex impacts on GNSS measurements, making the error distribution biased and multimodal. Future research can adopt biased GMM distribution with more components to establish a more accurate error model. Moreover, the AIPEM algorithm is built upon the assumption that the measurement set is dominated by healthy measurements. If the measurement set is dominated by unhealthy measurements, the adaptive error modeling process will be considerably affected. Future work should integrate additional sensors, such as the fisheye camera, to identify the number of unhealthy measurements and provide information on adjusting the error model. In addition, the initialization of error models is realized by goGPS weighting, which produces the same initial error profile in scenarios with different degrees of urbanization. To expedite the adaptation of error models to different environments, future research could employ informed initialization based on environmental context. One potential solution is to introduce environmental classifiers, such as machine learning models, to enhance the accuracy and efficiency of the initialization process.

APPENDIX A EFFECTS OF FAULT DETECTION AND EXCLUSION ON MEASUREMENT ERROR DISTRIBUTION

Due to the existence of unmodelled errors in measurements, the distribution of measurement error tends to be biased and has heavy tails. By applying fault detection and exclusion, we can improve the consistency between measurements and refine the bias and heavy tail issues in the error distribution. By adaptively modeling the measurement error after exclusion, we can capture the impacts of changing environments on the measurements, which also benefits the positioning solution.

In this appendix, we show the effects of fault detection and exclusion on measurement error distribution by conducting an additional experiment on the TST dataset. Specifically, we aim to examine the change in the tailedness and asymmetry of the error distribution before and after FDE. The kurtosis is commonly used to quantify the tailedness of a distribution

[44]. The kurtosis of the distribution of the random variable X is defined as the fourth standardized moment as follows:

$$\text{Kurt}(X) = E \left(\frac{X - \mu}{\sigma} \right)^4, \quad (25)$$

where μ is the mean of X , and σ is the standard deviation of X . A large kurtosis value indicates the existence of heavy tails. In addition, the skewness is commonly used to quantify the asymmetry of a distribution [44]. The kurtosis of the distribution of X is defined as the third standardized moment as follows:

$$\text{Skew}(X) = E \left(\frac{X - \mu}{\sigma} \right)^3. \quad (26)$$

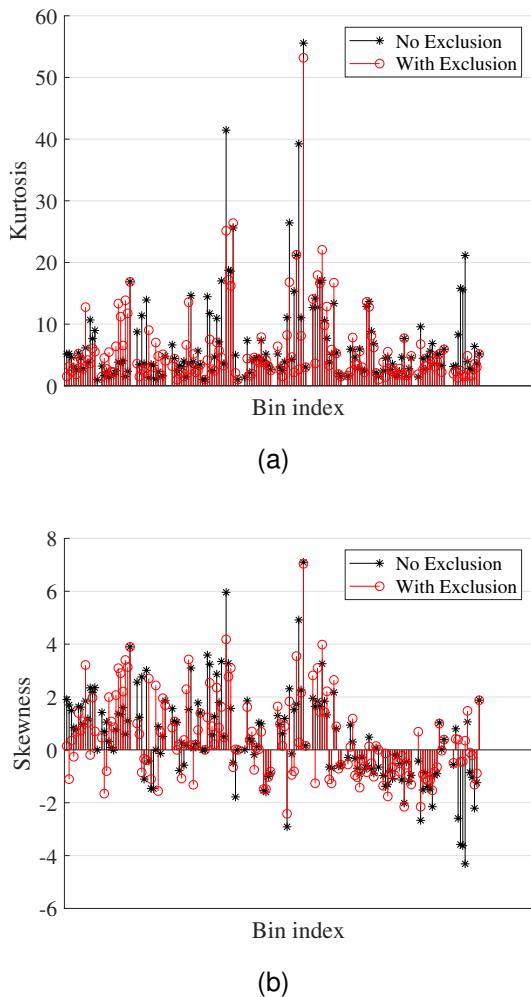


Fig. 9. (a) The kurtosis and (b) the skewness of the measurement error distribution in each bin.

A large magnitude of the skewness indicates the asymmetry of a distribution. In the TST dataset, we record the measurements before applying FDE (the raw measurement set) and the remaining measurements after FDE (the remaining measurement set) for each bin. We use the ground truth position of the receiver to calculate the measurement error in each set, based on which we can obtain the error distribution without and with FDE. We calculate the kurtosis and skewness of the error distribution without and with FDE and plot the

results in Fig. 9. As can be seen, the FDE process can reduce the kurtosis and the absolute value of skewness of error distribution nearly in all bins. Fig. 10(a) shows the error distribution in the Bin ($\lambda = 38 \sim 40 \text{ dB Hz}$, $\theta = 30 \sim 35^\circ$). It is obvious that the FDE process eliminates large error terms ($> 10 \text{ m}$) in the raw error distribution, which improves the consistency of measurements. In statistics, the kurtosis of the error distribution is reduced from 5.65 to 2.44 (a 56.81% reduction) while the skewness is reduced from 1.79 to 0.75 (a 58.10% reduction). Similar results are found in the Bin ($\lambda = 34 \sim 36 \text{ dB Hz}$, $\theta = 45 \sim 50^\circ$) in Fig. 10(b), where the kurtosis and skewness are reduced by 93.17% and 94.31%, respectively. Therefore, the FDE process can alleviate the tailedness and bias problem in the measurement error distribution in urban scenarios.

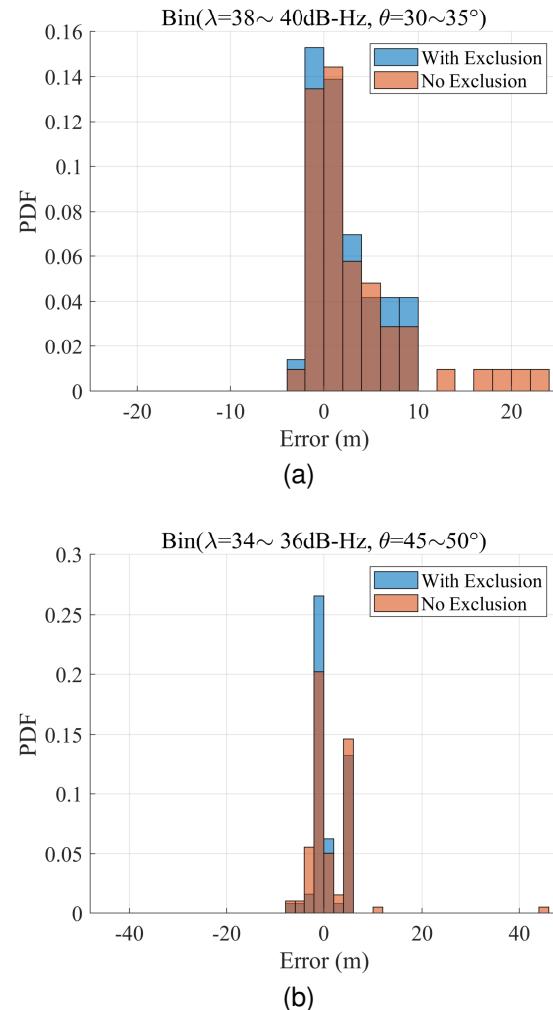


Fig. 10. (a) The distribution of measurement error in the Bin ($\lambda = 38 \sim 40 \text{ dB Hz}$, $\theta = 30 \sim 35^\circ$). When no exclusion is applied, the kurtosis is 5.65 and the skewness is 1.79; When exclusion is applied, the kurtosis is 2.44 and the skewness is 0.75. (b) The distribution of measurement error in the Bin ($\lambda = 34 \sim 36 \text{ dB Hz}$, $\theta = 45 \sim 50^\circ$). When no exclusion is applied, the kurtosis is 39.23 and the skewness is 4.92; When exclusion is applied, the kurtosis is 2.68 and the skewness is 0.28.

APPENDIX B COMPUTATION LOAD IN ERROR MODELING PROCESS

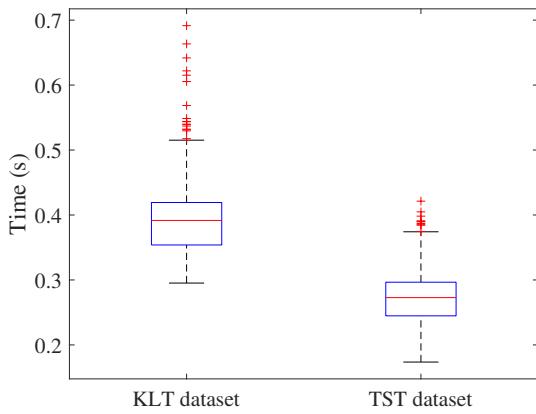


Fig. 11. The box plot of the computation time of the error modeling process in the KLT and TST datasets. All computations are conducted on a laptop (Intel Core i7-12700H CPU, 2.30 GHz).

REFERENCES

- [1] P. Misra and P. Enge, *Global positioning system: Signals, measurements, and performance*, 2nd ed. Lincoln, Mass: Ganga-Jamuna Press, 2006.
- [2] T. Ando, W. Kugimiyama, T. Hashimoto, F. Momiyama, K. Aoki, and K. Nakano, “Lateral control in precision docking using RTK-GNSS/INS and LiDAR for localization,” *IEEE Trans. Intell. Veh.*, vol. 6, no. 1, pp. 78–87, 2020.
- [3] J. Dahl, G. R. de Campos, C. Olsson, and J. Fredriksson, “Collision avoidance: A literature review on threat-assessment techniques,” *IEEE Trans. Intell. Veh.*, vol. 4, no. 1, pp. 101–113, 2018.
- [4] R. Toledo-Moreo and M. A. Zamora-Izquierdo, “Collision avoidance support in roads with lateral and longitudinal maneuver prediction by fusing GPS/IMU and digital maps,” *Transportation Research part C: emerging technologies*, vol. 18, no. 4, pp. 611–625, 2010.
- [5] B. Paden, M. Čáp, S. Z. Yong, D. Yershov, and E. Frazzoli, “A survey of motion planning and control techniques for self-driving urban vehicles,” *IEEE Trans. Intell. Veh.*, vol. 1, no. 1, pp. 33–55, 2016.
- [6] K. Chu, M. Lee, and M. Sunwoo, “Local path planning for off-road autonomous driving with avoidance of static obstacles,” *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1599–1616, 2012.
- [7] F. Pomerleau, P. Krüsi, F. Colas, P. Furgale, and R. Siegwart, “Long-term 3D map maintenance in dynamic environments,” in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 3712–3719.
- [8] C. Zhang, Z. Huang, M. H. Ang, and D. Rus, “Lidar degradation quantification for autonomous driving in rain,” in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 3458–3464.
- [9] J. R. V. Rivero, T. Gerbich, B. Buschardt, and J. Chen, “The effect of spray water on an automotive LiDAR sensor: A real-time simulation study,” *IEEE Trans. Intell. Veh.*, vol. 7, no. 1, pp. 57–72, 2021.
- [10] N. Zhu, J. Marais, D. Bétaillie, and M. Berbineau, “GNSS position integrity in urban environments: A review of literature,” *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 9, pp. 2762–2778, 2018.
- [11] L.-T. Hsu, “Analysis and modeling GPS NLOS effect in highly urbanized area,” *GPS Solut.*, vol. 22, no. 1, p. 7, 2018.
- [12] Z. Gao, S. X. Ding, and C. Cecati, “Real-time fault diagnosis and fault-tolerant control,” *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3752–3756, 2015.
- [13] J. Blanch, T. Walter, P. Enge, Y. Lee, B. Pervan, M. Rippl, and A. Spletter, “Advanced RAIM user algorithm description: Integrity support message processing, fault detection, exclusion, and protection level calculation,” in *Proc. 25th Int. Tech. Meeting Satell. Division Inst. Navigat.*, Nashville, TN, Sep. 2012, pp. 2828–2849.
- [14] B. W. Parkinson and P. Axelrad, “Autonomous GPS Integrity Monitoring Using the Pseudorange Residual,” *Navig. J. Inst. Navig.*, vol. 35, no. 2, pp. 255–274, Jun. 1988.
- [15] R. G. Brown, “A baseline GPS RAIM scheme and a note on the equivalence of three RAIM methods,” *Navig. J. Inst. Navig.*, vol. 39, no. 3, pp. 301–316, 1992.
- [16] M. Joerger and B. Pervan, “Fault detection and exclusion using solution separation and chi-squared ARAIM,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 2, pp. 726–742, Apr. 2016.
- [17] FAA, “Phase II of the GNSS evolutionary architecture study,” Federal Aviation Administration, Tech. Rep., 2010.
- [18] L.-T. Hsu, H. Tokura, N. Kubo, Y. Gu, and S. Kamijo, “Multiple faulty GNSS measurement exclusion based on consistency check in urban canyons,” *IEEE Sens. J.*, vol. 17, no. 6, pp. 1909–1917, 2017.
- [19] M. A. Sturza, “Navigation system integrity monitoring using redundant measurements,” *NAVIGATION*, vol. 35, no. 4, pp. 483–501, 1988.
- [20] T. Walter and P. Enge, “Weighted RAIM for precision approach,” in *Proc. 8th Int. Tech. Meeting Satell. Division Inst. Navigat.*, Palm Springs, CA, Sep. 1995, pp. 1995–2004.
- [21] R. G. Brown and P. W. McBurney, “Self-contained GPS integrity check using maximum solution separation,” *Navig. J. Inst. Navig.*, vol. 35, no. 1, pp. 41–53, 1988.
- [22] X. Niu, Q. Chen, Q. Zhang, H. Zhang, J. Niu, K. Chen, C. Shi, and J. Liu, “Using Allan variance to analyze the error characteristics of GNSS positioning,” *GPS Solutions*, vol. 18, no. 2, pp. 231–242, 2014.
- [23] R. Braff and C. Shively, “A method of over bounding ground based augmentation system (GBAS) heavy tail error distributions,” *J. Navig.*, vol. 58, no. 1, pp. 83–103, 2005.
- [24] P. Yan, W. Wen, F. Huang, and L.-T. Hsu, “A fault detection algorithm for lidar/imu integrated localization systems with non-gaussian noises,” in *Proceedings of the 2024 International Technical Meeting of The Institute of Navigation*, 2024, pp. 561–574.
- [25] E. Wang, C. Jia, G. Tong, P. Qu, X. Lan, and T. Pang, “Fault detection and isolation in GPS receiver autonomous integrity monitoring based on chaos particle swarm optimization-particle filter algorithm,” *Advances in Space Research*, vol. 61, no. 5, pp. 1260–1272, 2018.
- [26] P. He, G. Liu, C. Tan, and Y.-e. Lu, “Nonlinear fault detection threshold optimization method for RAIM algorithm using a heuristic approach,” *GPS Solutions*, vol. 20, pp. 863–875, 2016.
- [27] C. Zhang, C. Cao, C. Guo, T. Li, and M. Guo, “Navigation multisensor fault diagnosis approach for an unmanned surface vessel adopted particle-filter method,” *IEEE Sensors Journal*, vol. 21, no. 23, pp. 27 093–27 105, 2021.
- [28] J. Wang, C. Xu, M. Shi, and Z. Lu, “Protection level for precise point positioning based on Gaussian mixture model,” in *China Satellite Navigation Conference*. Springer, 2022, pp. 45–55.
- [29] Y. Yun, H. Yun, D. Kim, and C. Kee, “A Gaussian sum filter approach for DGNSS integrity monitoring,” *The Journal of Navigation*, vol. 61, no. 4, pp. 687–703, 2008.
- [30] J. Blanch, T. Walker, P. Enge, Y. Lee, B. Pervan, M. Rippl, A. Spletter, and V. Kropp, “Baseline advanced RAIM user algorithm and possible improvements,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 1, pp. 713–732, 2015.
- [31] A. M. Herrera, H. F. Suhandri, E. Realini, M. Reguzzoni, and M. C. de Lacy, “goGPS: open-source MATLAB software,” *GPS Solut.*, vol. 20, pp. 595–603, 2016.
- [32] X. Liu, W. Wen, and L.-T. Hsu, “GLIO: Tightly-coupled GNSS/LiDAR/IMU integration for continuous and drift-free state estimation of intelligent vehicles in urban areas,” *IEEE Trans. Intell. Veh.*, 2023.
- [33] A. Chalvatzaras, I. Pratikakis, and A. A. Amanatiadis, “A survey on map-based localization techniques for autonomous vehicles,” *IEEE Trans. Intell. Veh.*, vol. 8, no. 2, pp. 1574–1596, 2022.
- [34] T. Pfeifer and P. Protzel, “Expectation-maximization for adaptive mixture models in graph optimization,” in *2019 international conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 3151–3157.
- [35] D. M. Rosen, M. Kaess, and J. J. Leonard, “Robust incremental online inference over sparse factor graphs: Beyond the Gaussian case,” in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 1025–1032.
- [36] T. Takasu and A. Yasuda, “Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB,” in *Proc. Int. Symp. GPS/GNSS*, vol. 1. International Convention Center Jeju Korea Seogwipo-si, Republic of Korea, 2009, pp. 1–6.
- [37] A. El-Mowafy, B. Xu, and L.-T. Hsu, “Integrity monitoring using multi-GNSS pseudorange observations in the urban environment combining ARAIM and 3D city models,” *Journal of Spatial Science*, vol. 67, no. 1, pp. 91–110, 2022.

- [38] X. Xia, L.-T. Hsu, and W. Wen, "Integrity-constrained factor graph optimization for GNSS positioning," in *Proc. IEEE/ION Position Location Navigation Symp.* IEEE, 2023, pp. 414–420.
- [39] T. Walter, K. Gunning, R. Eric Phelts, and J. Blanch, "Validation of the unfaulted error bounds for ARAIM," *Navig. J. Inst. Navig.*, vol. 65, no. 1, pp. 117–133, 2018.
- [40] J. Blanch, T. Walter, and P. Enge, "Position error bound calculation for GNSS using measurement residuals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 977–984, Jul. 2008.
- [41] Y. Yun, H. Yun, D. Kim, and C. Kee, "A Gaussian sum filter approach for DGNSS integrity monitoring," *J. Navig.*, vol. 61, no. 4, pp. 687–703, 2008.
- [42] J. Lee, "Laas position domain monitor analysis and test results for CAT II/III operations," in *Proc. 17th Int. Tech. Meeting Satell. Division Inst. Navigat.*, Long Beach, CA, Sep. 2004, pp. 2786–2796.
- [43] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. B*, vol. 39, no. 1, pp. 1–22, Sep. 1977.
- [44] N. Weiss, P. Holmes, and M. Hardy, *A Course in Probability*. Pearson Addison Wesley, 2006.
- [45] L.-T. Hsu, F. Huang, H.-F. Ng, G. Zhang, Y. Zhong, X. Bai, and W. Wen, "Hong Kong UrbanNav: An open-source multisensory dataset for benchmarking urban navigation algorithms," *Navig. J. Inst. Navig.*, vol. 70, no. 4, 2023.



Penggao Yan received the bachelor's degree in Communication Engineering in 2018 and the master's degree in Pattern Recognition and Intelligent Systems in 2021, both from Wuhan University, China. He is currently a Ph.D. Candidate at the Department of Aeronautical and Aviation Engineering, Faculty of Engineering, Hong Kong Polytechnic University.

His research interests include non-Gaussian noise modeling, fault detection and integrity monitoring in localization systems, and control-aided localization.



Xiao Xia received his B.Sc. degree in Vehicle Engineering from Beijing Institute of Technology, China in 2018, and M.Sc. degree in Automation Engineering from RWTH, Aachen University, Germany in 2021. He is currently pursuing a Ph.D. degree at the Department of Aeronautical and Aviation Engineering, Faculty of Engineering, Hong Kong Polytechnic University.

His research interests include integrity monitoring for safety-certified localization solutions, variational inference, robust estimation, and constrained optimization.



Michele Brizzi (Member, IEEE) is a post-doc researcher at the Department of Industrial, Electronic and Mechanical Engineering, Roma Tre University. He received the B.Sc. in Electronic Engineering, the M.Sc. in Information and Communication Technology Engineering, and his Ph.D. in Applied Electronics from Roma Tre University in 2016, 2018, and 2022, respectively.

His main research interests are in the field of computer vision, sensor fusion, and navigation algorithms for autonomous vehicles.



Weisong Wen (Member, IEEE) received a BEng degree in Mechanical Engineering from Beijing Information Science and Technology University (BISTU), Beijing, China, in 2015, and an MEng degree in Mechanical Engineering from the China Agricultural University, in 2017. After that, he received a Ph.D. degree in mechanical engineering, the Hong Kong Polytechnic University. He was a visiting student researcher at the University of California, Berkeley (UCB) in 2018. He is currently an assistant professor in the Department of Aeronautical and Aviation Engineering, the Hong Kong Polytechnic University.

His research interests include multi-sensor integrated localization for autonomous vehicles, SLAM, and GNSS positioning in urban canyons.



Li-Ta Hsu (Senior Member, IEEE) received the B.S. and Ph.D. degrees in aeronautics and astronautics from National Cheng Kung University, Taiwan, in 2007 and 2013, respectively. He is currently an associate professor at the Department of Aeronautical and Aviation Engineering, Faculty of Engineering, Hong Kong Polytechnic University, before he served as a post-doctoral researcher in Institute of Industrial Science at University of Tokyo, Japan. In 2012, he was a visiting scholar in University College London, U.K. He is an Associate Fellow of RIN.

His research interests include GNSS positioning in challenging environments and localization for pedestrians, autonomous driving vehicles, and unmanned aerial vehicles.