Performance of ambiguity-resolved detector for GNSS mixed-integer model

Yin, Chengyu; Teunissen, P. J.G.; Tiberius, C. C.J.M.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Performance of ambiguity-resolved detector for GNSS mixed-integer model

Chengyu Yin[1] · P. J. G. Teunissen[1,2,3] · C. C. J. M. Tiberius[1]

Abstract

Teunissen (J Geod 98(83):1–16, 2024) proposed the ambiguity-resolved (AR) detection theory for GNSS mixed-integer model validation. In this contribution, we study the performance of the AR detector through analysis and simulation experiments and compare it with the ambiguity-float (AF) and ambiguity-known (AK) detectors. We describe how the detectors can be implemented and how to evaluate their performance by computing the power as functions of the model misspecifications' size. We present two simulation experiments with single- and dual-frequency GPS models and demonstrate that the AR detector can provide a larger detection power than the AF detector, even if the success rate is not close to one. Then, we obtain power functions over 25 user locations with five observation models and 72 satellite geometries per location per model. We find that the AR detector increases the detection probability of ionosphere and troposphere delays by 47% and 60% on average when the success rate is larger than 97.5% and the level of significance is 0.01. We also find the AR detection power to be larger than that of the AF detector in case of multi-dimensional misspecifications.

Keywords GNSS · Mixed-integer model · Ambiguity resolution · Model validation · Monte Carlo simulation

Introduction

The Global Navigation Satellite System (GNSS) observation models connect the observables and the unknown model parameters. Unmodeled effects may misspecify the assumed model, which can seriously deteriorate the estimation results when they remain unnoticed (Misra and Enge 2011; Hofmann-Wellenhof et al. 2012; Teunissen and Montenbruck 2017). Therefore, the validation of the observation model is essential in GNSS data processing. The detection, identification, and adaptation (DIA) procedure has been widely used in model validation (Baarda 1968; Kok 1984; Teunissen 2000). Detection is the first step in this procedure, where an overall model test is performed to diagnose if a model misspecification occurs.

Currently, the detection of GNSS mixed-integer model misspecifications is either based on the ambiguity-float (AF) detector or the ambiguity-known (AK) detector (Teunissen 2024). The AF detector is applied when the ambiguity is not resolved (fixed) to an integer vector. The ambiguity is considered to be an *unknown real* vector, and knowledge of its integer property is then not used to benefit detection. The AK detector, on the other hand, can be applied when the ambiguity is completely *known*. In practice, ambiguity resolution is carried out when the success rate is very close to one (e.g., >0.999), and the resolved ambiguity is commonly assumed to be known and consequently treated as a deterministic quantity (Feng et al. 2009; Wang et al. 2022; Zhang et al. 2023), although it is a random estimator. Teunissen (2024) proposes the ambiguity-resolved (AR) detection theory, in which the ambiguity is treated as an *unknown integer* vector. By considering the distributional property of the resolved ambiguity, the AR-detector enables ambiguity resolution to contribute to model validation even if the success rate is not close to one. The distribution of the AR test statistic cannot be written in a closed form due to the discrete nature of the resolved ambiguity. Numerical simulations and analysis are necessary to understand the

✉ Chengyu Yin
  c.yin@tudelft.nl

1 Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands

2 Department of Infrastructure Engineering, The University of Melbourne, Melbourne, Australia

3 GNSS Research Centre, Curtin University, Perth, Australia

performance of the AR detector for detecting GNSS model misspecifications.

We evaluate and compare the performance of the AF, AK, and AR detectors for the single-differenced (SD) common-oscillator model and the short baseline double-differenced (DD) model. The SD model with a common oscillator is used in attitude determination, which can be affected by constant biases (Keong 1999; Chen 2016). The short baseline DD model is one of the commonly used relative positioning models, in which the pseudorange and carrier phase observables from more than one receiver are combined to eliminate or reduce the common errors between the receivers (Leick et al. 2015; Odijk 2017). The relative positioning models can be formulated with undifferenced (UD), SD, or DD observables collected from a single baseline or a network (Tiberius 1998; Odijk 2017). We use the short baseline DD model as an example to conduct experiments and compare the performance of the detectors. Similar performance can be expected for the models formulated with UD and SD observables, as well as the atmosphere-corrected network-RTK and PPP-RTK user models, which are intrinsically equivalent to the short-baseline DD model given that the user observation setup is the same and the variance-covariance (vc) matrix of the corrections are appropriately considered (Teunissen and Khodabandeh 2015; Odijk et al. 2016; Kouba et al. 2017). We focus on the validation of relative positioning models, which are vulnerable to blunders in pseudorange observables, carrier phase outliers due to multipath or faulty corrections and cycle slips (Braasch 2017; Duan et al. 2024), and ionosphere and troposphere delays that are not fully canceled or corrected (Ahn et al. 2006; Lawrence et al. 2006; Huang and van Graas 2007; Wanninger 2004; Hernández-Pajares et al. 2011; Dao et al. 2024). These misspecifications are modeled by means of the alternative hypotheses H_a, for which the detection power will be analyzed.

This contribution starts with the theoretical background: the observation models under the null and alternative hypotheses are described, and the ambiguity resolution and misspecification detection theory for the GNSS mixed-integer models is reviewed. After that, we describe how to carry out the mixed-integer model detection and how to obtain the statistical power of the detectors. The critical value and power of the AR detector can only be obtained by the Monte Carlo method; thus, we introduce how to evaluate the uncertainty of the Monte Carlo simulations. It is followed by the experiment section that compares the power of detectors for one- and multi-dimensional model misspecifications, where the superiority of the AR detector to the AF detector is shown, and the irregular shapes of the AR power functions are explained. Finally, the summary and conclusions are presented.

Review of theory

Differential observation model

Assume $s + 1$ satellites from the same constellation are tracked by the rover and reference receivers on f frequencies for k epochs. The linearized double-differenced (DD) GNSS observation model can be written as

$$
\begin{aligned}
H_0 : & \quad E\{\underline{y}\} = Aa + Bb, \; D\{\underline{y}\} = Q_{yy}, \\
H_a : & \quad E\{\underline{y}\} = Aa + Bb + Cc, \; D\{\underline{y}\} = Q_{yy},
\end{aligned}
\tag{1}
$$

where H_0 and H_a denote the null and alternative hypotheses, respectively. $E\{\cdot\}$ is the expectation operator and $D\{\cdot\}$ is the dispersion operator. The underline $'\underline{\cdot}'$ denotes a random variable or vector. $\underline{y} = [\underline{\varphi}^T, \underline{p}^T]^T \in \mathbb{R}^{2sfk}$ contains the double-differenced carrier phase $\underline{\varphi}$ and pseudorange $\underline{p}$ and we assume them to be normally distributed. $\mathbb{R}^{(*)}$ represents a $(*)$ dimensional real space. With this definition of H_a, we consider only misspecifications on the mean of $\underline{y}$, which are modeled by design matrix C and vector $c \in \mathbb{R}^q$, with q the dimension of the misspecification under H_a. $a \in \mathbb{Z}^{sf}$ is the unknown integer ambiguity vector and $b \in \mathbb{R}^3$ is the real-valued baseline vector. $\mathbb{Z}^{(*)}$ represents a $(*)$ dimensional integer space. The design matrix $[A, B, C]$ is assumed to be of full column rank.

The design matrix for the ambiguity vector is

$$
A = \begin{bmatrix} 1 & 0 \end{bmatrix}^T \otimes \mathbf{1}_k \otimes \mathrm{diag}\left(\lambda_1, \cdots, \lambda_f\right) \otimes I_s,
\tag{2}
$$

where $\otimes$ denotes the Kronecker product, $\mathbf{1}_k$ is a $k \times 1$ vector with values of 1, 'diag' refers to a diagonal matrix, λ_f is the wavelength of the f-th frequency, and I_s is an $s \times s$ identity matrix.

The design matrix for the baseline is

$$
B = \mathbf{1}_2 \otimes M_k \otimes \mathbf{1}_f \otimes \left(D^T G\right),
\tag{3}
$$

with $M_k = \mathbf{1}_k$ for a stationary receiver and $M_k = I_k$ for a moving receiver; $D^T = [-\mathbf{1}_s, I_s]$ the $s \times (s + 1)$ between-satellite differencing matrix ($D^T = I_{s+1}$ for the SD model in which the between-satellite differencing is not conducted); G the $(s + 1) \times 3$ geometry matrix contains the unit vectors from satellites to receivers. In this contribution, we consider the instantaneous and short time span cases where G is practically time-invariant.

The variance-covariance (vc) matrix of observable $\underline{y}$ is

$$
Q_{yy} = \mathrm{diag}\left(\sigma_\varphi^2, \sigma_p^2\right) \otimes R_k \otimes I_f \otimes 2D^T W^{-1} D,
\tag{4}
$$

σ_φ^2 and σ_p^2 are the zenith-referenced undifferenced phase and pseudorange variances; we assume that the phase and pseudorange observables are uncorrelated. The non-diagonal elements of the $k \times k$ matrix R_k account for the time correlation between the observables from different epochs, and its diagonal values are 1 (Odijk and Teunissen 2008). With the $f \times f$ identity matrix I_f, we assume the observables from different frequencies are uncorrelated and of the same precision. $W = \text{diag}(w_1, \cdots, w_i, \cdots, w_{s+1})$ contains the elevation-dependent weights for each satellite (Euler and Goad 1991),

$$w_i = 1/\left(a_0 + a_1 \exp\left(-E^i/E_0\right)\right)^2, \tag{5}$$

where E^i is the elevation of the i-th satellite in degrees. Examples of the model coefficients a_0, a_1 and E_0 can be found in (Euler and Goad 1991; Jin and de Jong 1996). $a_0 = 1$, $a_1 = 10$ and $E_0 = 10$ degrees are used in our simulation experiments.

When the DD model contains observables from two constellations, the design matrices and vc-matrix can be stacked as follows,

$$A^{12} = \begin{bmatrix} A^1 & \\ & A^2 \end{bmatrix}, B^{12} = \begin{bmatrix} B^1 \\ B^2 \end{bmatrix}, Q_{yy}^{12} = \begin{bmatrix} Q_{yy}^1 & \\ & Q_{yy}^2 \end{bmatrix}, \tag{6}$$

where A^i, B^i and Q_{yy}^i are the design matrices and vc-matrix for the i-th constellation. The observables are differenced within each constellation, and the receiver hardware delays are canceled; thus, inter-system biases do not appear in the model.

To evaluate the performance of the detectors for a specific model misspecification, we consider the observation model under H_a, in which C models a misspecification of size c. We evaluate the detection performance of four misspecifications with the DD model and of one misspecification with the SD common oscillator model. The design matrices for the one-dimensional misspecification are given below, and they can be concatenated when we work with multi-dimensional misspecifications, e.g., blunders in several pseudorange observables.

Although most errors and delays are eliminated or corrected in relative positioning models, pseudorange multipath delays may appear and affect the estimation accuracy, ranging from several meters to even 100 m in most severe conditions (Braasch 2017). The design matrix for a blunder in pseudorange observable of the i-th satellite, j-th frequency, k-th epoch is

$$C = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes c^k \otimes c^j \otimes D^T c^i, \tag{7}$$

where c^i is a canonical vector with its i-th entry equal to one and zero for others.

Phase observables may contain outliers and cycle slips. Additionally, phase bias corrections in the network-based scenario may contain errors that can be interpreted as phase outliers in the user model at the centimeters level (e.g., Duan et al. 2024). To evaluate the performance of the detectors for phase outliers, we use the model with two epochs, and we assume the outlier or cycle slip appears in the observables of the second epoch. For the phase outlier from the i-th satellite, j-th frequency, the corresponding design matrix is written as

$$C = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes c^j \otimes D^T c^i. \tag{8}$$

The troposphere delay can be separated into dry and wet constituents, of which the dry component can be modeled with high precision. In relative positioning models, the wet delay could be canceled or corrected in most cases. However, during severe weather conditions, troposphere anomalies can cause delays ranging from several centimeters to several decimeters in the differenced/corrected observables (Ahn et al. 2006; Lawrence et al. 2006; Huang and Van Graas 2007). The design matrix for the troposphere delay at the k-th epoch is

$$C = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes c^k \otimes \mathbf{1}_f \otimes D^T \begin{bmatrix} m\left(E^1\right) & \cdots & m\left(E^{s+1}\right) \end{bmatrix}^T, \tag{9}$$

with $m(\cdot)$ the troposphere mapping function. $m(E) = 1/\sin(E)$ is used in our simulations (Hobinger and Jakowski 2017). The misspecification c in this model is the differential zenith wet delay in meters.

Similar to the troposphere delay, the differential ionosphere delay can be ignored under the nominal condition, while this is not the case under anomalous conditions, for example, in the presence of ionosphere disturbance (Wanninger 2004; Hernández-Pajares et al. 2011) or when the ionosphere spatial gradients are hundreds of millimeters per kilometer, which can lead to differential ionosphere delays at the decimeter level (Pullen et al. 2009). The performance of the AR detector for ionosphere delays of satellite i on epoch k is evaluated with the design matrix

$$C = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \otimes c^k \otimes \mu \otimes D^T c^i, \tag{10}$$

where $\mu = \begin{bmatrix} \frac{\lambda_1^2}{\lambda_1^2} & \cdots & \frac{\lambda_f^2}{\lambda_1^2} \end{bmatrix}^T$. The misspecification c in this model is the size of the differential ionosphere delay of the first frequency in meters. The value can be

converted to TECU by multiplying the delay (in meters) by $\frac{2f_1^2}{80.6 \times 10^{16}} \approx 6.1587 \text{ TECU/m}$.

The between-receiver single differenced model with a common oscillator is used for attitude determination (Chen 2016). This model can be affected by constant biases, including the initial bias and the line bias caused by the different lengths and materials of the cables. Constant biases are at the centimeter level and are common for all signals (Keong 1999); thus, all entries of the design matrix are one,

$$C = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \mathbf{1}_k \otimes \mathbf{1}_f \otimes \mathbf{1}_{s+1}. \tag{11}$$

The SD functional and stochastic models have the same structures as the DD model, where the between-satellite difference matrix D^T is replaced by an identity matrix.

Ambiguity resolution

Ambiguity resolution is the key step in GNSS model validation with the AR detector (Teunissen 2024), which is a mapping from a float ambiguity estimator to a resolved ambiguity estimator, denoted as $\breve{a} = \mathcal{I}(\hat{a})$. Three classes of integer ambiguity estimators have been developed (Teunissen 2017): integer estimator, integer aperture estimator, and integer equivariant estimator. In this contribution, we restrict ourselves to the class of integer estimators (Teunissen 1999a), where the ambiguity is always resolved to an integer vector; thus $\breve{a} \in \mathbb{Z}^n$, with n the dimension of the ambiguity vector.

The resolved integer ambiguity $\breve{a}$ is obtained in two steps. First, the integer property of the ambiguity vector is ignored. The float estimator $\hat{a}$ and its vc-matrix $Q_{\hat{a}\hat{a}}$ are

$$\hat{a} = \bar{A}^+ y, \quad Q_{\hat{a}\hat{a}} = \left(\bar{A}^T Q_{yy}^{-1} \bar{A}\right)^{-1}, \tag{12}$$

where $\bar{A}^+ = \left(\bar{A}^T Q_{yy}^{-1} \bar{A}\right)^{-1} \bar{A}^T Q_{yy}^{-1}$, and $\bar{A} = P_B^\perp A$ with $P_B^\perp = I - B\left(B^T Q_{yy}^{-1} B\right)^{-1} B^T Q_{yy}^{-1}$. The float ambiguity is resolved to $\breve{a}$ in the second step with one of the integer estimators (Teunissen 1999a), in which the integer least-squares (ILS) estimator has the largest success rate of fixing the float ambiguity to the correct integer (Teunissen 1999b) and thus provides performance closest to the AK detector. The ILS solution can be obtained efficiently by the LAMBDA method (Teunissen 1995), which conducts the integer search based on the decorrelation-transformed ambiguity vector $\hat{z}$ and its vc-matrix $Q_{\hat{z}\hat{z}}$,

$$\hat{z} = Z^T \hat{a}, \quad Q_{\hat{z}\hat{z}} = Z^T Q_{\hat{a}\hat{a}} Z, \tag{13}$$

where Z^T is an admissible ambiguity decorrelation transformation matrix (ibid.).

The success rate of the ILS estimator cannot be computed analytically, and the integer bootstrapping (IB) success rate for the decorrelated ambiguity $\hat{z}$ provides a tight and easy-to-compute lower bound for it (Teunissen 1998; Verhagen 2003), which is computed as

$$P(\breve{z}_{IB} = Z^T a) = \prod_{i=1}^n \left[2\Phi\left(\frac{1}{2\sigma_{\hat{z}_{i|1,\cdots,i-1}}}\right) - 1 \right], \tag{14}$$

where a is the true ambiguity; $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}v^2\right\} dv$ is the cumulative distribution function (CDF) of the standard normal distribution; the conditional standard deviations $\sigma_{\hat{z}_{i|1,\cdots,i-1}}$ are the square root of diagonal values of the diagonal matrix provided by the triangular matrix factorization of $Q_{\hat{a}\hat{a}}$.

Detection theory

Two detectors that are currently used to detect GNSS model misspecifications are the AF and AK detectors. The AF detector ignores the integer property of the ambiguity and is based on the ambiguity-float residual

$$\hat{e} = P_{[A,B]}^\perp y. \tag{15}$$

The AK detector employs the ambiguity-known residual computed with the known vector a and can only be used when the ambiguity is fully known,

$$\hat{e}(a) = P_B^\perp (y - Aa). \tag{16}$$

By replacing the known a in (16) with the resolved integer ambiguity $\breve{a}$, we obtain the ambiguity-resolved residual,

$$\breve{e} = \hat{e}(\breve{a}) = P_B^\perp (y - A\breve{a}). \tag{17}$$

Test statistics of the AF, AK, and AR detectors are the corresponding squared norms of the residuals. The AF and AK test statistics and their distributions under H_0 are written as,

$$\begin{aligned} AF &: \|\hat{e}\|_{Q_{yy}}^2 |H_0 \sim \chi^2(r,0), \\ AK &: \|\hat{e}(a)\|_{Q_{yy}}^2 |H_0 \sim \chi^2(r(a),0), \end{aligned} \tag{18}$$

where r is the redundancy of the model under H_0 and $r(a)$ is the redundancy of the model when the known ambiguity is excluded from the unknown vector. $\|v\|_M^2 = v^T M^{-1} v$ denotes the squared norm of v weighted by M^{-1}. The AR detector test statistic can be written as (Teunissen 2024)

$$AR : \|\check{\underline{e}}\|^2_{Q_{yy}} = \|\hat{\underline{e}}\|^2_{Q_{yy}} + \|\check{\underline{e}}\|^2_{Q_{\hat{a}\hat{a}}}, \tag{19}$$

where $\check{\underline{e}} = \hat{\underline{a}} - \check{\underline{a}}$ is the ambiguity residual. The probability density function (PDF) of $\check{\underline{e}}$ is written as (Teunissen 2002)

$$f_{\check{\underline{e}}}(x) = \frac{\sum_{z \in \mathbb{Z}^n} \exp\left\{-\frac{1}{2}\|x + z\|^2_{Q_{\hat{a}\hat{a}}}\right\}}{\sqrt{|2\pi Q_{\hat{a}\hat{a}}|}} s_0(x), \tag{20}$$

with $s_0(x)$ the indicator function of the integer estimators' pull-in region of zero vector (Teunissen 2017). With these test statistics, the detectors then read.

$$
\begin{aligned}
AF &: \text{Reject } H_0 \text{ if } \|\hat{e}\|^2_{Q_{yy}} > \chi^2_\alpha(r, 0), \\
AK &: \text{Reject } H_0 \text{ if } \|\hat{e}(a)\|^2_{Q_{yy}} > \chi^2_\alpha(r(a), 0), \\
AR &: \text{Reject if } \|\check{e}\|^2_{Q_{yy}} > \kappa_\alpha,
\end{aligned}
\tag{21}
$$

where α is the given level of significance and $\chi^2_\alpha(r, 0)$ refers to the $1 - \alpha$ quantile of the central χ^2 PDF with r degrees of freedom. The AR critical value κ_α cannot be obtained analytically due to the irregular distribution of the AR test statistic; the method to obtain κ_α will be introduced in the following section.

The model under H_a is used to evaluate the performance of the detectors for a specific model misspecification, modeled by the additional term Cc. The distributions of the AF and AK test statistics under H_a are written as

$$
\begin{aligned}
AF &: \|\hat{\underline{e}}\|^2_{Q_{yy}}|H_a \sim \chi^2(r, \lambda_{\hat{e}}), \\
AK &: \|\hat{\underline{e}}(a)\|^2_{Q_{yy}}|H_a \sim \chi^2(r(a), \lambda_{\hat{e}(a)}),
\end{aligned}
\tag{22}
$$

with the noncentrality parameters

$$\lambda_{\hat{e}} = \|P^\perp_{[A,B]}Cc\|^2_{Q_{yy}}, \quad \lambda_{\hat{e}(a)} = \|P^\perp_B Cc\|^2_{Q_{yy}}.$$

The distribution of AR test statistic under H_a cannot be written in a closed form; an example is shown in Fig. 1.

Performance evaluation

Implementation of detectors

The critical values of the AF and AK detectors can be obtained with the CDF of the central chi-squared distribution with degrees of freedom of r and $r(a)$, denoted as $\chi^2_\alpha(r, 0)$, and $\chi^2_\alpha(r(a), 0)$, respectively. The critical value of the AR detector κ_α fulfills

$$\int_0^{\kappa_\alpha} f_{\|\check{\underline{e}}\|^2_{Q_{yy}}|H_0}(x)\, dx = 1 - \alpha. \tag{23}$$

Due to the complexity of the distribution of $\|\check{\underline{e}}\|^2_{Q_{yy}}$, we obtain κ_α through Monte Carlo simulation (Metropolis and Ulam 1949). $\|\check{\underline{e}}\|^2_{Q_{yy}}$ can be represented as a sum of $\|\hat{\underline{e}}\|^2_{Q_{yy}}$ and the $\|\check{\underline{e}}\|^2_{Q_{\hat{a}\hat{a}}}$ (19). Since $\|\hat{\underline{e}}\|^2_{Q_{yy}}$ and $\|\check{\underline{e}}\|^2_{Q_{\hat{a}\hat{a}}}$ are independent, we can first generate samples of them separately and then obtain samples of $\|\check{\underline{e}}\|^2_{Q_{yy}}$. According to the 'remove-restore' property of the integer estimator (Teunissen 1999b), samples of the ambiguity residual can be obtained without knowing the exact value of the true integer ambiguity. We define the number of samples to be generated as N_c and carry out the Monte Carlo simulation using the following steps.

1) Generate N_c samples of the float ambiguity vector, $\hat{a}_0^{(1)}, \cdots, \hat{a}_0^{(i)}, \cdots, \hat{a}_0^{(N_c)}$, that follow a normal distribution according to $\mathcal{N}_n(0, Q_{\hat{a}\hat{a}})$, which denotes an n-dimensional normal distribution with zero mean and variance-covariance matrix $Q_{\hat{a}\hat{a}}$.
2) Conduct ambiguity resolution (Teunissen 1999a) to obtain samples of the resolved ambiguity, $\check{a}_0^{(1)}, \cdots, \check{a}_0^{(i)}, \cdots, \check{a}_0^{(N_c)}$. Then samples of $\|\check{\underline{e}}\|^2_{Q_{\hat{a}\hat{a}}}$ can be obtained with $\|\check{e}_0\|^{2,(i)}_{Q_{\hat{a}\hat{a}}} = \|\hat{a}_0^{(i)} - \check{a}_0^{(i)}\|^2_{Q_{\hat{a}\hat{a}}}$.
3) Generate N_c samples of the AF residual, $\|\hat{e}_0\|^{2,(1)}_{Q_{yy}}, \cdots, \|\hat{e}_0\|^{2,(i)}_{Q_{yy}}, \cdots, \|\hat{e}_0\|^{2,(N_c)}_{Q_{yy}}$, that follow a $\chi^2(r, 0)$ distribution.
4) Obtain samples of the AR test statistic with $\|\check{e}_0\|^{2,(i)}_{Q_{yy}} = \|\hat{e}_0\|^{2,(i)}_{Q_{yy}} + \|\check{e}_0\|^{2,(i)}_{Q_{\hat{a}\hat{a}}}$.
5) Finally, sort the samples $\|\check{e}_0\|^{2,(i)}_{Q_{yy}}$ in ascending order, the critical value is approximated by the $[(1 - \alpha)N_c]$-th ordered sample, $[\cdot]$ denotes the rounding operation.

The number of samples N_c determines the precision of the critical value simulation. It can be chosen according to Yin et al. (2024), where the relation between N_c and simulation uncertainty is investigated.

Uncertainty of the AR critical value

The procedure to obtain the AR critical value falls under the topic of Monte Carlo quantile simulation, as the unknown critical value is the $1 - \alpha$ quantile of the PDF $f_{\|\check{\underline{e}}\|^2_{Q_{yy}}|H_0}(x)$, and it is approximated by the $[(1 - \alpha)N_c]$-th ordered sample. The distributional property of the ordered sample and the quantile simulated with the Monte Carlo method are introduced by Serfling (1980), where several approaches to evaluate the uncertainty of the quantile simulation are
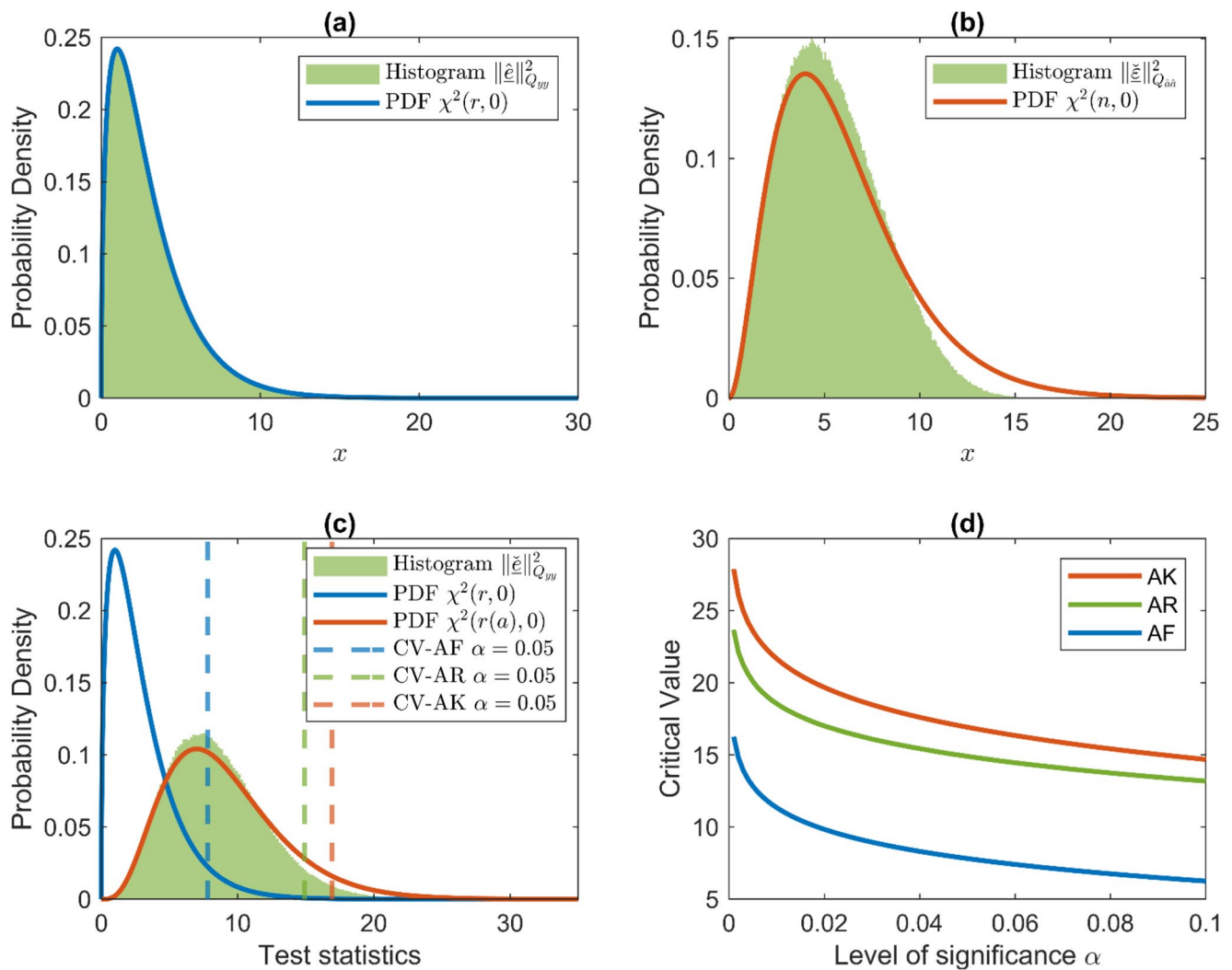
Fig. 1 (**a**) Normalized histogram of $\|\hat{e}\|^2_{Q_{yy}}$ samples under H_0, and $\chi^2(r,0)$ PDF. (**b**) Normalized histogram of $\|\breve{\varepsilon}\|^2_{Q_{\hat{a}\hat{a}}}$ samples under H_0, and $\chi^2(n,0)$ PDF. (**c**) Normalized histogram of AR test statis- tics and PDFs of AR and AK test statistics; dashed lines give critical values of detectors for $\alpha = 0.05$. (**d**) Critical values of detectors for different α

also provided. We describe one of the approaches based on asymptotic normality.

It is shown by Serfling (1980) that the simulated $1 - \alpha$ quantile $\hat{\underline{\kappa}}_\alpha$ with N_c samples follows

$$\sqrt{N_c}\left(\hat{\underline{\kappa}}_\alpha - \kappa_\alpha\right) \rightarrow \mathcal{N}_1\left(0, \sigma^2\right), \text{ when } N_c \rightarrow \infty,$$
$$\text{with } \sigma^2 = \frac{\alpha\left(1 - \alpha\right)}{f^2_{\|\breve{\varepsilon}\|^2_{Q_{yy}}|H_0}\left(\kappa_\alpha\right)} \tag{24}$$

in which κ_α denotes the true but unknown critical value. This property provides the asymptotic variance σ^2 and indicates that $\hat{\underline{\kappa}}_\alpha$ is asymptotically consistent. Although $f_{\|\breve{\varepsilon}\|^2_{Q_{yy}}|H_0}\left(\kappa_\alpha\right)$ is unknown, it can be approximated by the probability density around $\hat{\kappa}_\alpha$ obtained based on the samples. Then σ^2/N_c approximates the variance of the

simulated critical value. A detailed example of this method can be found in (Yin et al. 2024).

Statistical power of detectors

We evaluate the performance of the AF, AK, and AR detectors in the following way:

- We first select a misspecification, which is modeled by the term Cc under H_a.
- Then, we obtain the statistical power of the AK, AF, and AR detector to detect the selected misspecification for a particular size c and level of significance α;
- We compare the performance of the detectors by comparing the power.

Under H_a, AF and AK test statistics follow non-central chi-squared distributions (22), and the power to detect a selected model misspecification with size c can be obtained with the CDF of their distributions as

$$
\begin{aligned}
\gamma_{AF} &= P\left[\|\hat{\underline{e}}\|_{Q_{yy}}^2 > \chi_\alpha^2(r,0)|H_a\right], \\
\gamma_{AK} &= P\left[\|\hat{\underline{e}}(a)\|_{Q_{yy}}^2 > \chi_\alpha^2(r(a),0)|H_a\right].
\end{aligned}
\tag{25}
$$

We obtain now the power of the ambiguity-resolved detector by Monte Carlo simulation. The mean of $\hat{\underline{a}}$ under H_a is

$$
E\{\hat{\underline{a}}|H_a\} = a + \bar{A}^+ Cc
\tag{26}
$$

and the noncentrality parameter of $\|\hat{\underline{e}}\|_{Q_{yy}}^2$ is computed according to (22). Similar to the critical value simulation, we generate samples of $\|\hat{\underline{e}}\|_{Q_{yy}}^2$ and $\|\check{\underline{\varepsilon}}\|_{Q_{\hat{a}\hat{a}}}^2$ independently with $a = 0$. We simulate the power using N_p samples as follows.

1) Generate N_p samples of the float ambiguity vector $\hat{a}_a^{(1)}, \cdots, \hat{a}_a^{(i)}, \cdots, \hat{a}_a^{(N_p)}$, which follow a normal distribution according to $\mathcal{N}_n(E\{\hat{\underline{a}}|H_a\}, Q_{\hat{a}\hat{a}})$.
2) Conduct ambiguity resolution (Teunissen [1999a]) to obtain samples of the resolved ambiguity $\check{a}_a^{(1)}, \cdots, \check{a}_a^{(i)}, \cdots, \check{a}_a^{(N_p)}$. Then samples of $\|\check{\underline{\varepsilon}}\|_{Q_{\hat{a}\hat{a}}}^2$ can be obtained with $\|\check{\varepsilon}_a\|_{Q_{\hat{a}\hat{a}}}^{2,(i)} = \|\hat{a}_a^{(i)} - \check{a}_a^{(i)}\|_{Q_{\hat{a}\hat{a}}}^2$.
3) Generate N_p samples of the AF residual $\|\hat{e}_a\|_{Q_{yy}}^{2,(1)}, \cdots, \|\hat{e}_a\|_{Q_{yy}}^{2,(i)}, \cdots, \|\hat{e}_a\|_{Q_{yy}}^{2,(N_p)}$, that follow a $\chi^2(r, \lambda_{\hat{e}})$ distribution (22).
4) Samples of the AR test statistic can be computed as $\|\check{e}_a\|_{Q_{yy}}^{2,(i)} = \|\hat{e}_a\|_{Q_{yy}}^{2,(i)} + \|\check{\varepsilon}_a\|_{Q_{\hat{a}\hat{a}}}^{2,(i)}$.
5) Finally, count the number of samples $\|\check{e}_a\|_{Q_{yy}}^{2,(i)}$ larger than the critical value, denoted as n_p. The simulated power in percentage follows as $\hat{\gamma}(c) = 100 \times \frac{n_p}{N_p}\%$.

Uncertainty of AR detection power

The power of the AR detector is obtained by a two-stage Monte Carlo simulation procedure. We evaluate the uncertainty of the power simulation empirically by repeating the simulation for N_r times (Morio and Balesdent [2015]).

1) Assume we use a total number of N_1 samples to simulate the critical value and N_2 samples to simulate the power. We conduct N_r repeated simulations, and each time with $N_c = N_1/N_r$ for the critical value and $N_p = N_2/N_r$ for the power. N_r can be chosen as 50 or

100 in practice (Morio and Balesdent [2015]; El Masri et al. [2021]).
2) Simulate the critical value and the power for the selected misspecification with size c. Repeat this step N_r times and get N_r power simulations denoted as $\hat{\gamma}_i(c), i = 1, 2, \cdots, N_r$.
3) The mean of N_r power simulations is the final simulated power,

$$
\hat{\gamma}(c) = \frac{1}{N_r}\sum_{i=1}^{N_r}\hat{\gamma}_i(c).
\tag{27}
$$

Its standard deviation is computed assuming $\hat{\gamma}_i(c)$ are independent and of the same precision,

$$
\sigma_{\hat{\gamma}(c)} = \sqrt{\frac{1}{N_r}\sigma_{\hat{\gamma}_i(c)}}, \text{ with } \sigma_{\hat{\gamma}_i(c)} = \sqrt{\frac{\sum_{i=1}^{N_r}(\hat{\gamma}_i(c) - \hat{\gamma}(c))^2}{N_r - 1}}.
\tag{28}
$$

Experiments

In this section, simulation experiments are conducted to evaluate the performance of the AR detector. We first consider one-dimensional misspecifications. An experiment with a single-frequency GPS model is carried out, where the distributions of the test statistics under H_0 and H_a and the detection power of detectors as functions of misspecifications' size are presented. Then, the power functions for a dual-frequency GPS model are provided. After that, the performance of the AR detector is evaluated by simulation experiments over 25 user locations and five models with 72 satellite geometries per location per model. Finally, we present three examples of detecting multi-dimensional misspecifications. The experiments for phase outlier detection employ two-epoch models, and single-epoch models are used for the other misspecifications.

Performance for single-frequency model

In this simulation experiment, we consider single-frequency (L1) observables from GPS satellites. Seven satellites are observed, and the skyplot is shown in Fig. 2. To compute Q_{yy} with (4), the zenith-referenced standard deviations for pseudorange and carrier phase observables are set as $\sigma_p = 0.2$m and $\sigma_\varphi = 0.002$m, respectively.

PDFs under H_0

We evaluate the performance of detecting a blunder in pseudorange, a phase outlier, and atmosphere delays with the DD observation model. The ILS success rate (14) of the
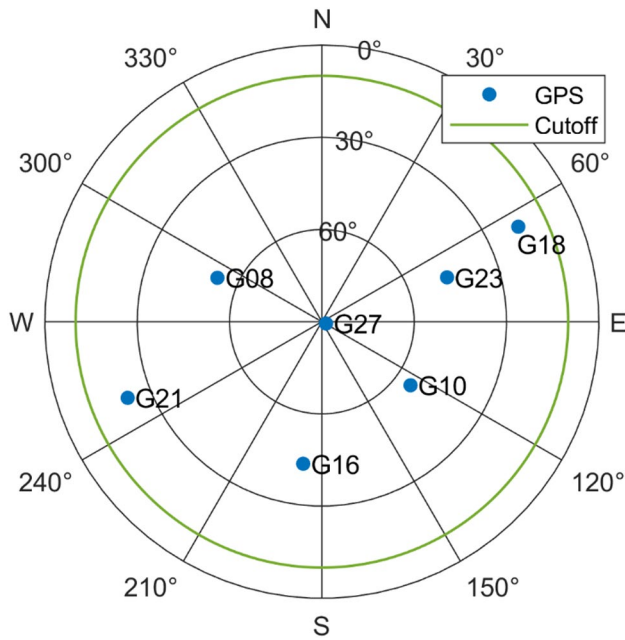
Fig. 2 Skyplot of GPS satellites for single-frequency experiment

single-epoch ambiguity resolution is 80.5%. As we describe in 'Implementation of detectors', the critical value of the AR detector is obtained by Monte Carlo simulation. The histograms in Fig. 1 are normalized to provide the probability density. We use 10^6 samples for Monte Carlo simulation with 200 bins in each histogram. Figure 1a presents the distribution $\chi^2(r,0)$ and normalized histogram of $\|\hat{\underline{e}}\|^2_{Q_{yy}}$ samples under H_0. Figure 1b compares the normalized histogram of $\|\check{\underline{\varepsilon}}\|^2_{Q_{\hat{a}\hat{a}}}$ samples under H_0, and the PDF of $\chi^2(n,0)$, which is the distribution of $\|\hat{\underline{e}}(a)\|^2_{Q_{\hat{a}\hat{a}}}$ with the n dimensional known ambiguity vector. $\check{\underline{\varepsilon}}$ is always bounded inside the pull-in region of the integer estimator; thus, its square norm is also bounded. As a result, the normalized histogram of $\|\check{\underline{\varepsilon}}\|^2_{Q_{\hat{a}\hat{a}}}$ samples is compressed along the horizontal axis compared with $\chi^2(n,0)$. Fig. 1c exhibits the distributions of three detectors. The AF and AK test statistics follow chi-squared distributions (18), and the distribution of the AR test statistic lies between them. The reason is that the AR test statistic is compressed compared with the AK test statistic, as shown in Fig. 1b, and is larger than the AF test statistic due to including the square norm of ambiguity residual. Figure 1d presents the critical values of the three detectors for different levels of significance.
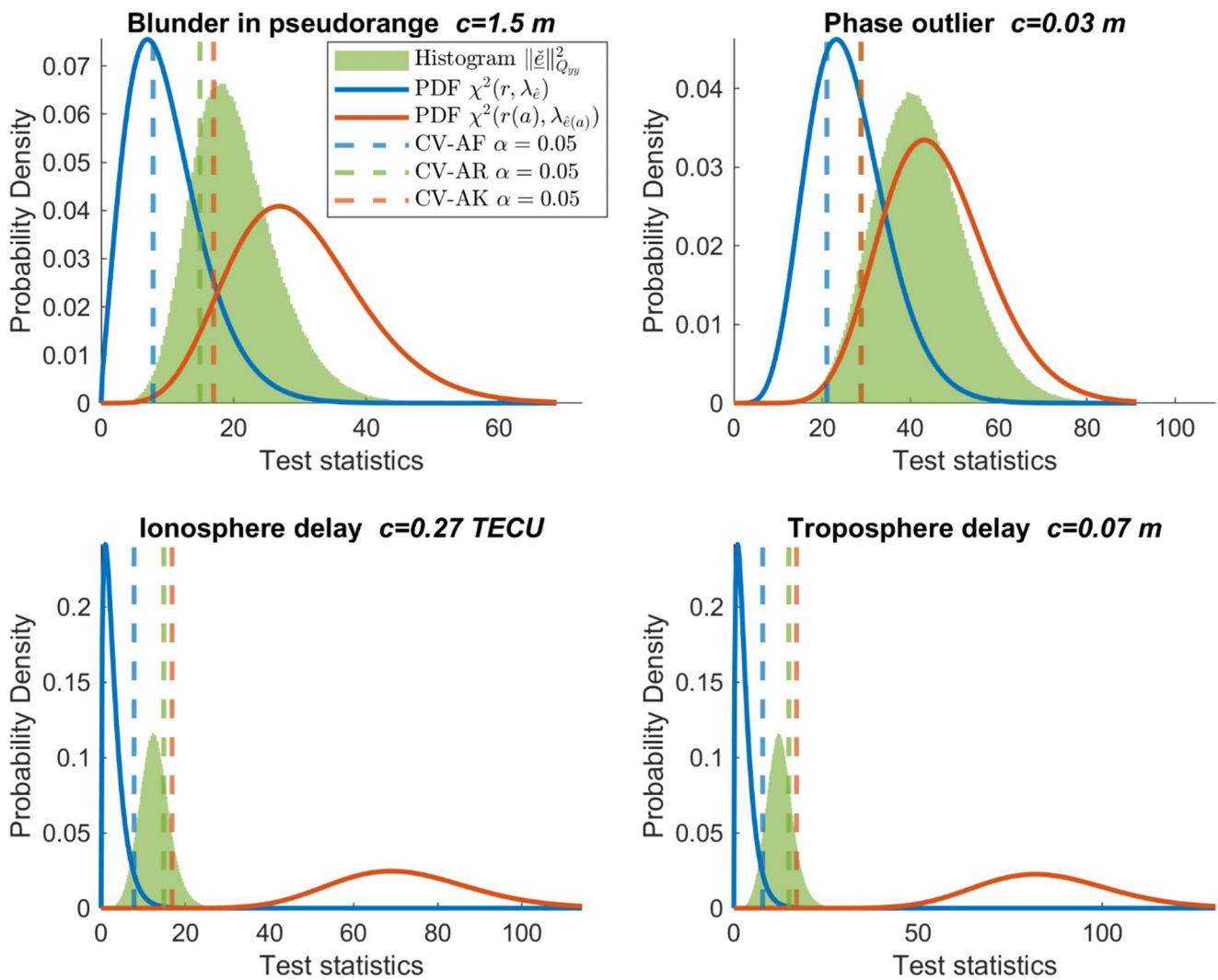
PDFs under H_a

Figure 3 shows the distributions of the test statistics under H_a for four one-dimensional misspecifications, and the corresponding detection powers are given in Table 1. Compared with the distributions under H_0 (Fig. 1c), the high-density regions of the distributions under H_a are shifted toward the larger values. This is due to the noncentrality introduced by the misspecifications to the AF and AK test statistics. For the AR test statistic, the change in its distribution combines the noncentrality of $\|\hat{\underline{e}}\|^2_{Q_{yy}}|H_a$ and the bias of $\hat{\underline{a}}|H_a$ that affects the distribution of $\|\check{\underline{\varepsilon}}\|^2_{Q_{\hat{a}\hat{a}}}|H_a$. The impact of $\|\check{\underline{\varepsilon}}\|^2_{Q_{\hat{a}\hat{a}}}$ will be further explained in Fig. 4. The detection power is the probability that a misspecification (with size c) can be detected. For a blunder in the pseudorange of satellite G08 with the size of 1.5 m, the AR detector increases the power by 20%. The phase outlier of satellite G08 with 3 cm can be detected with a 92% power by the AR detector, which is higher than the 68% float detection power. The AF detector can hardly detect the differential atmosphere delays of several centimeters in this experiment. However, once the ambiguity is known, the AK detector can detect them with a power equal to one. This can be explained by the closed form of the AF and AK noncentrality parameters (Eqs. 22 and 23 in Teunissen 2024). Once the ambiguity is known, the noncentrality parameters of the ionosphere and troposphere delays are magnified by the factor σ_p^2/σ_ϕ^2, which equals 10,000 in this experiment and drives the difference in the detection power of AF and AK detectors. Although the single-epoch DD ambiguity resolution success rate in this experiment is only 80.5%, we still observe the improvement in the detection power of the AR detector compared with the AF detector. Table 1 also shows the significant differences between the power of the AR and AK detectors. Since the ambiguities are unknown in practice even when they are integer-estimated, one should not predict the performance of the AR detector relying on the assumed probabilistic property of the AK detector.

Power functions

Figure 5 presents the power of the AF, AK, and AR detectors as functions of four misspecifications' sizes evaluated with the single frequency model. The vertical error bars with $\pm 2\sigma$ for the simulated AR powers are plotted for the blunder in pseudorange and phase outlier power functions. The power and uncertainty are obtained with the steps described in 'Uncertainty of AR detection power' with $N_1 = N_2 = 2 \times 10^5$ and $N_r = 50$. The powers are evaluated for 31 c values for the blunder in pseudorange and phase outlier, and 201 c values for the atmosphere delays.

The AF and AK power functions are smooth curves since the noncentrality of their test statistics monotonically increases with c. However, as observed in the atmosphere delay experiments, the AR power function can be

Fig. 3 Distributions of three test statistics under H_a for four one-dimensional misspecifications. Blue and red lines are distributions of AF and AK test statistics. Normalized histogram in green illustrates AR distribution. Dashed lines give their critical values for $\alpha = 0.05$

Table 1 Detection power corresponding to Fig. 3 ($\alpha = 0.05$)

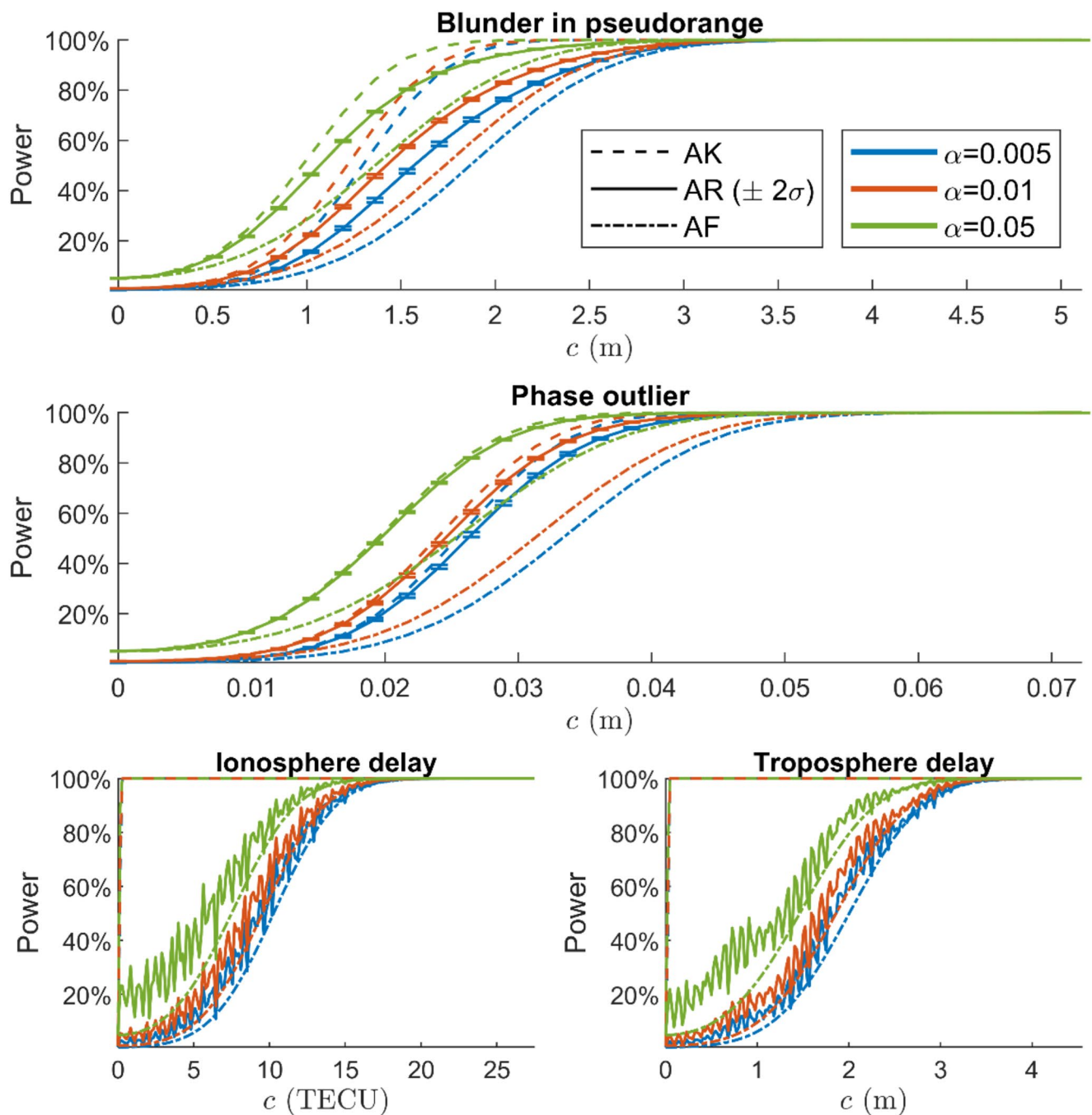| Misspecification	Size (c)	AF	AR	AK
Blunder in pseudorange (G08)	1.5 m	59%	79%	92%
Phase outlier (G08)	0.03 m	68%	92%	100%
Ionosphere delay (G08)	0.27 TECU	5%	26%	100%
Troposphere delay	0.07 m	5%	24%	100%

non-monotonic, which can be explained as follows. The AR power can be written as (Teunissen 2024)

$$\gamma_{AR} = \int P\left[\|\check{\varepsilon}\|_{Q_{\hat{a}\hat{a}}}^2 > \kappa_\alpha - x | H_a\right] f_{\underline{x}}(x | H_a)\, dx, \qquad (29)$$

where $\underline{x} = \|\hat{\underline{e}}\|_{Q_{yy}}^2$. With this formulation, the impact of Cc on $\|\hat{\underline{e}}\|_{Q_{yy}}^2$ is captured by $f_{\underline{x}}(x | H_a)$ and its impact on $\|\check{\varepsilon}\|_{Q_{\hat{a}\hat{a}}}^2$ is captured by the probability inside the
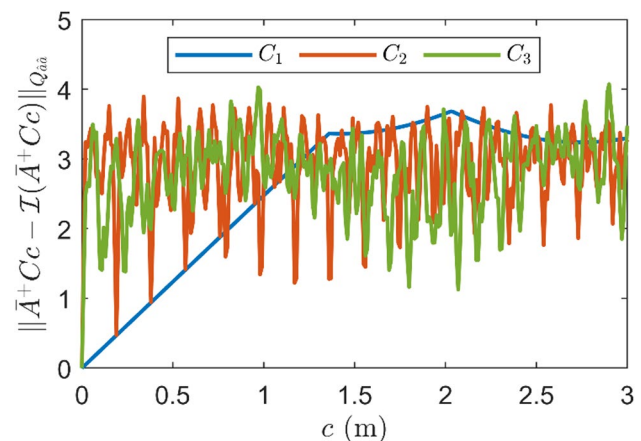
Fig. 4 Distance (under metric $Q_{\hat{a}\hat{a}}$) between $\bar{A}^+ Cc$ and nearest integer of it, denoted by $\mathcal{I}(\bar{A}^+ Cc)$, as functions of misspecifications' size for a blunder in pseudorange (C_1), ionosphere delay (C_2), and troposphere delay (C_3)

Fig. 5 Detection power for blunder in pseudorange (G08), phase outlier (G08), ionosphere (G08), and troposphere as functions of misspecifications' size with single-frequency model. The uncertainty of simulated AR power is shown by $\pm 2\sigma$ vertical error bars in first and second graphs

integral. Although the noncentrality of $\|\hat{\underline{e}}\|^2_{Q_{yy}}$ monotonically increases with c, this is not the case for the probability $P[\|\check{\underline{\varepsilon}}\|^2_{Q_{\hat{a}\hat{a}}} > \kappa_\alpha - x | H_a]$. This probability will be small if the bias of $\hat{\underline{a}} | H_a$, which equals $\bar{A}^+ Cc$, is close to an integer, and it will be larger otherwise. Figure 4 presents the distance between $\bar{A}^+ Cc$ and the nearest integer as functions of c for the blunder in pseudorange (C_1), ionosphere delay (C_2), and troposphere delay (C_3). For atmosphere delays,

the distance oscillates with c, while it increases smoothly for the pseudorange. The explanation for this behaviour is provided by Fig. 8 in (Teunissen 2024).

In almost all cases, the AR power functions lie between the AK and AF power functions. This means the AR detector provides a larger detection power than the AF detector for detecting the same misspecification. However, the AR detector may provide a lower power than the AF detector in an exceptional case where the bias $\bar{A}^+ Cc$ is close to an

integer. In this case, the additional term in the AR test statistic $\|\underaccent{\check{}}{\varepsilon}\|^2_{Q_{\hat{a}\hat{a}}}$ turns out to be small and does not contribute to the detection of the misspecification. This case rarely happens for several c values, as is shown in the power functions of the atmosphere delays in Fig. 5.

The power functions of the constant bias detection for the SD common oscillator model are presented in Fig. 7. The powers are evaluated for 101 c values. The ambiguity resolution success rate for this experiment is 99.3%, which is higher than that of the DD ambiguity since the between-satellite difference is not conducted. The receiver clock offsets and hardware delays are assumed to be canceled by the between-receiver differencing in the model under H_0. The detection power is evaluated for c in the range of 0 to 0.1 m since the constant bias is several centimeters, as shown by Keong (1999). The constant bias belongs to the 'tropo-type' misspecifications described by Teunissen (2024), which affects the pseudorange and the carrier phase identically. The noncentrality parameter of the AK detector is σ^2_p/σ^2_ϕ times larger than that of the AF detector. As is shown by the power functions, the AK detection power is close to 100% for c larger than 4 cm, while the AF detector is insensitive to the centimeter-level bias. The AR detection power is very close to that of the AK detector for the bias smaller than 4 cm. When bias increases, we observe a non-monotonic behavior that is driven by the distribution of $\|\underaccent{\check{}}{\varepsilon}\|^2_{Q_{\hat{a}\hat{a}}}$.

Performance for dual-frequency model

In this simulation experiment, we evaluate the performance of the AR detector with a dual-frequency (L1/L5) GPS observation model, and the skyplot of the satellites is shown in Fig. 8. The zenith-referenced standard deviations for pseudorange and carrier phase observables are set as $\sigma_p = 0.5$m and $\sigma_\varphi = 0.005$m, with which we assume the observables are collected from a low-cost receiver.

Figure 6 presents the power of the AF, AK, and AR detectors as functions of four misspecifications' size for the dual-frequency DD model; the ILS success rate of the single-epoch DD ambiguity resolution is 97.4%. Figure 9 shows the power functions of the constant bias detection for an SD common oscillator model; the ILS success rate of the SD ambiguity resolution is 99.9%. The number of samples for the Monte Carlo simulation and the number of c values used to compute powers are the same as the previous experiment with the single-frequency model. For detecting the blunder in pseudorange, phase outlier, and constant bias, the AR detector performs almost the same as the AK detector. Similar to Fig. 5, the power functions of the atmosphere delays are oscillating. Compared with the single-frequency experiments, the AR power functions in the dual-frequency

case are closer to the AK detector due to the higher success rate.

To provide a quantitative description of the improvement in the AR detection power compared with the AF detector, we compute the *average power difference* between the AR and AF detection power with

$$\frac{1}{N_d}\sum_{i=1}^{N_d}\gamma_{AR}\left(c_i\right) - \gamma_{AF}\left(c_i\right) \tag{30}$$

where N_d is the number of power points that are taken into account, c_i is the i-th c value for which powers are evaluated. We consider the powers fulfilling $10\% < \gamma_{AR}(c_i) < 90\%$, since the power functions are convergent when the power is close to α and 1 (See Table 2).

As is shown in Table 2, improvement of the power is limited for detecting the blunder in pseudorange, although the AR detector performs almost the same as the AK detector. This can be explained by the noncentrality parameter for 'code-type' misspecifications derived by Teunissen (2024), which shows that the noncentrality parameters of both the AK and AF detectors are driven by pseudorange precision. For detecting the phase outlier, we observe 28% and 26% increases in power on average for $\alpha = 0.005$ and $\alpha = 0.01$. The power increases are over 41% and up to 73% on average for detecting the atmosphere delays and constant bias. In this simulation experiment with the dual-frequency GPS observation model, the AR detector performs better than the float detector as it provides higher detection powers.

## Performance evaluation at different locations

We conduct simulation experiments over 25 user locations, which are shown in Fig. 10, with five types of observation models and obtain the power functions of the AR and AF detectors for four misspecifications: blunder in pseudorange, phase outlier, ionosphere, and troposphere delays. For each user location and model, we formulate the DD observation equations with 72 satellite geometries, obtained every 20 min for 24 h from the precise orbit product of the IGS (International GNSS Service).

We are interested in the performance of the AR detector when the ambiguity resolution success rate of the model is not close to one. Therefore, we conduct experiments with relatively weak models listed in Table 3. We obtain 7200 power functions for detecting a blunder in pseudorange and atmosphere delays and 5400 power functions for phase outlier detection. The average power difference is computed with (30) for all power functions to compare the performance of the AR detector with the AF detector.
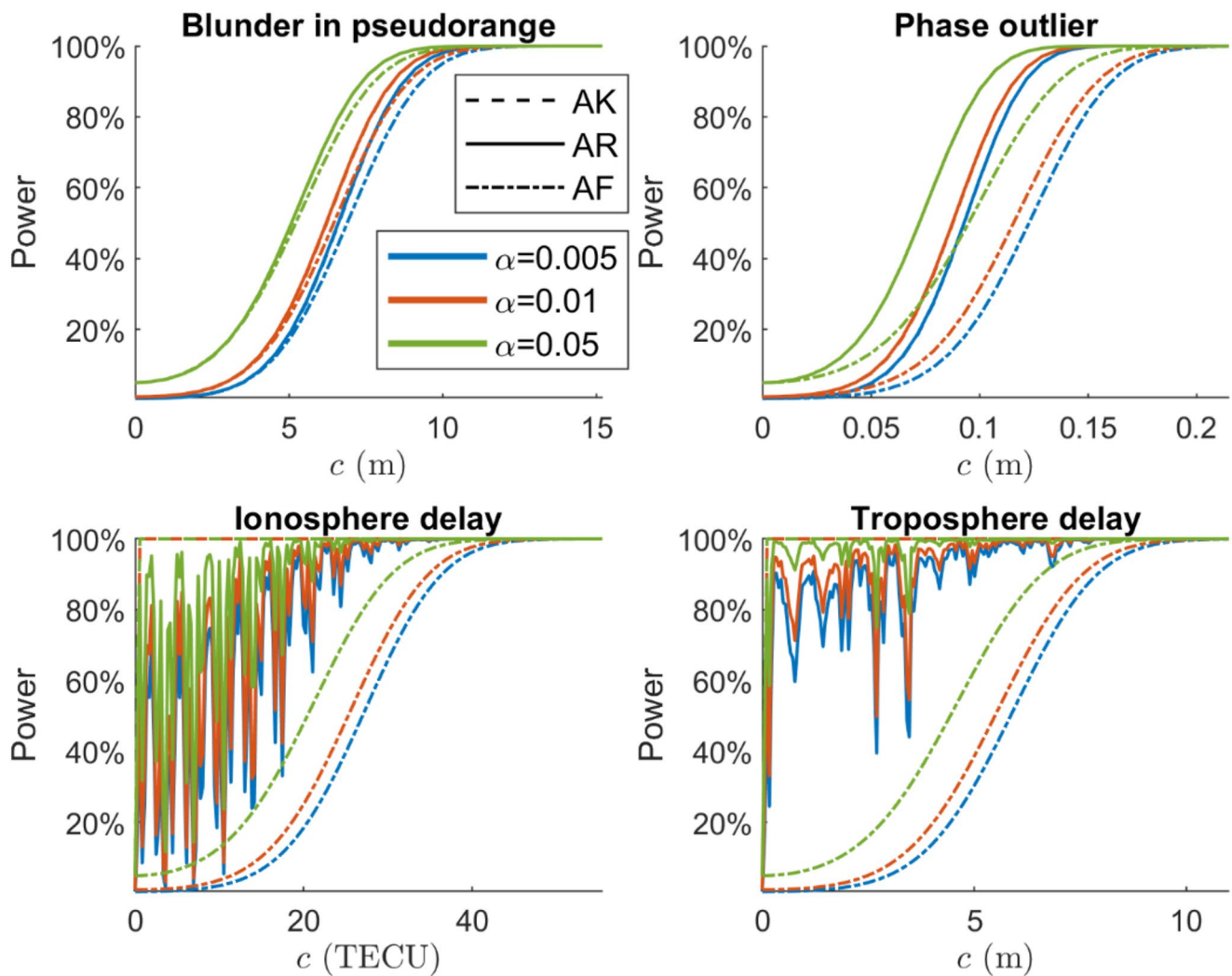
Fig. 6 Power functions of detecting blunder in pseudorange (G05), phase outlier (G05), ionosphere (G05) and troposphere delays with dual-frequency GPS model
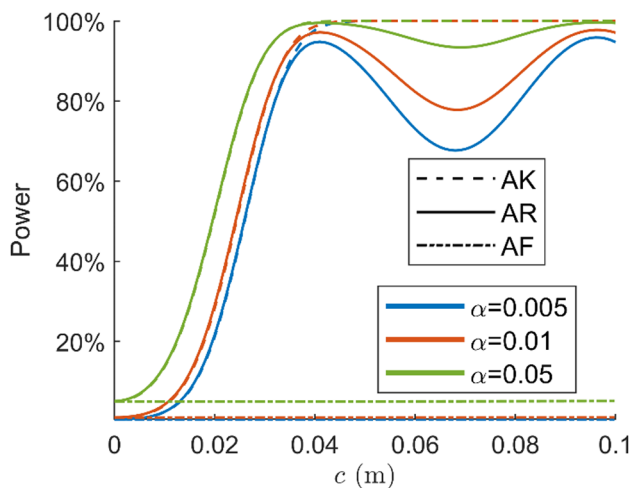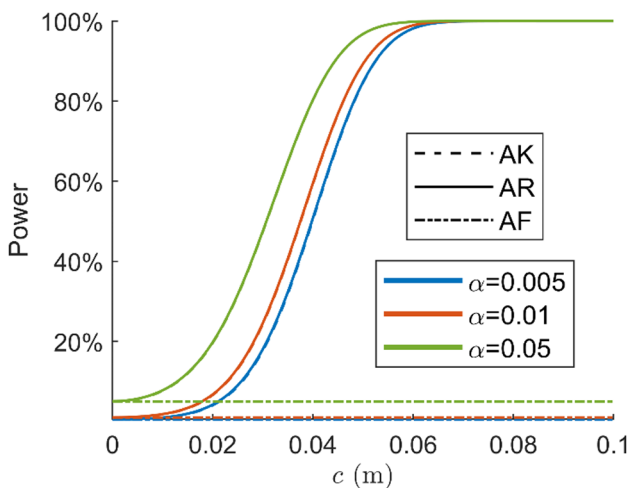

Fig. 7 Power functions of constant bias detection for single-frequency SD model

Figure 11 (left) shows the average power differences obtained from the power functions with $\alpha = 0.01$ for four misspecifications. Figure 11 (right) is obtained by splitting the success rate range $[0.6, 1]$ into 16 intervals and compute the overall average of power difference within each interval, four of which are provided in Table 4. Compared with the AF detector, the AR detector performs better in detecting atmospheric delays and phase outliers. Improvement can be observed even though the ambiguity resolution success rate is not close to one. For the experiments with success rates in $[0.775, 0.8]$, the detection powers are increased by 10% and 17% on average for ionosphere and troposphere delay detection. The improvement is significant when the success rate is close to one. For the experiments with success rates larger than 97.5%, the detection powers are increased by 12%, 47%, and 60% on average for phase outlier, ionosphere, and troposphere delay detection, respectively.
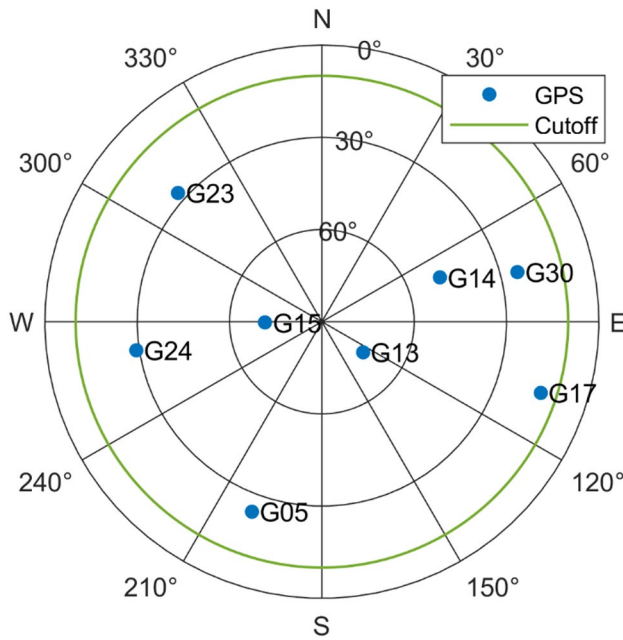
Fig. 8 Skyplot of GPS satellites for dual-frequency experiment

Fig. 9 Power functions of SD constant bias detection with dual-frequency model

Table 2 Average power difference of dual-frequency experiment

Misspecification	$\alpha = 0.005$	$\alpha = 0.01$	$\alpha = 0.05$
Blunder in pseudorange	5%	4%	2%
Phase outlier	28%	26%	19%
Ionosphere delay	49%	52%	51%
Troposphere delay	72%	73%	57%
Constant Bias	48%	47%	41%

The results show that the improvement by the AR detector for the atmosphere delay is larger than that for the blunder in pseudorange, and the AR detector can perform poorer than the AF detector in the latter case. Teunissen (2024) explains the difference between the AF and AK detectors for detecting different types of misspecifications, which also helps

**Fig. 10** User locations for AR detector performance evaluation

Table 3 Observation models for AR detector performance evaluation

Misspecifications	Constellation	Signal	Epochs	$\sigma_p(\sigma_\varphi = \sigma_p/100)$
Blunder in pseudorange	GPS	L1	1	{0.2, 0.35}m
Ionosphere delay	GPS	L1+L5	1	0.5 m
Troposphere delay	GPS+Galileo	L1+E1	1	0.5 m
Phase outlier	GPS	L1	2	{0.2, 0.35}m
	GPS+Galileo	L1+E1	2	0.5 m

to understand the performance of the AR detector since it lies between the AF and AK detectors. Although the AK test statistic has a larger noncentrality parameter, $\lambda_{\hat{e}(a)} > \lambda_{\hat{e}}$, it also has a larger critical value due to the stronger variability of $\|\underline{\hat{e}}(a)\|^2_{Q_{yy}}$. As a result, the AK detector can perform poorer than the AF detector if they have similar noncentrality parameters. For the ionosphere and troposphere delay detection, Teunissen (2024) shows that $\lambda_{\hat{e}(a)}$ is magnified by the factor σ_p^2/σ_ϕ^2, which explains the superiority of the AR detector. For the blunder in pseudorange detection, according to Eq. (20) in Teunissen (2024), the difference between $\lambda_{\hat{e}(a)}$ and $\lambda_{\hat{e}}$ is driven by the projection of $C_p c$ in the range space of the B_p, where C_p and B_p refer to the submatrices of C and B containing rows for the pseudorange observables. With the principal angle (Björck and Golub 1973; Teunissen 1997) between the range space of C_p and B_p, denoted as $\theta_{C_p B_p}$, the AK noncentrality parameter is magnified by a factor $\frac{1}{sin^2\left(\theta_{C_p B_p}\right)}$ for one-dimensional C. The DD model in the previous dual-frequency experiment can be taken as an example where $\theta_{C_p B_p} = 55.2^\circ$ and $\lambda_{\hat{e}(a)}/\lambda_{\hat{e}} = 1.48$. This factor is close to the ratio between the AK and AF critical values, which equals 1.79 for $\alpha = 0.01$ in this example. For the blunder in pseudorange detection, the change in the noncentrality parameters $\lambda_{\hat{e}(a)}$ and $\lambda_{\hat{e}}$ is not significant compared with the change in critical values. Hence, the AK detector performs similarly to the AF detector for detecting the blunder in pseudorange, and so does the AR detector.
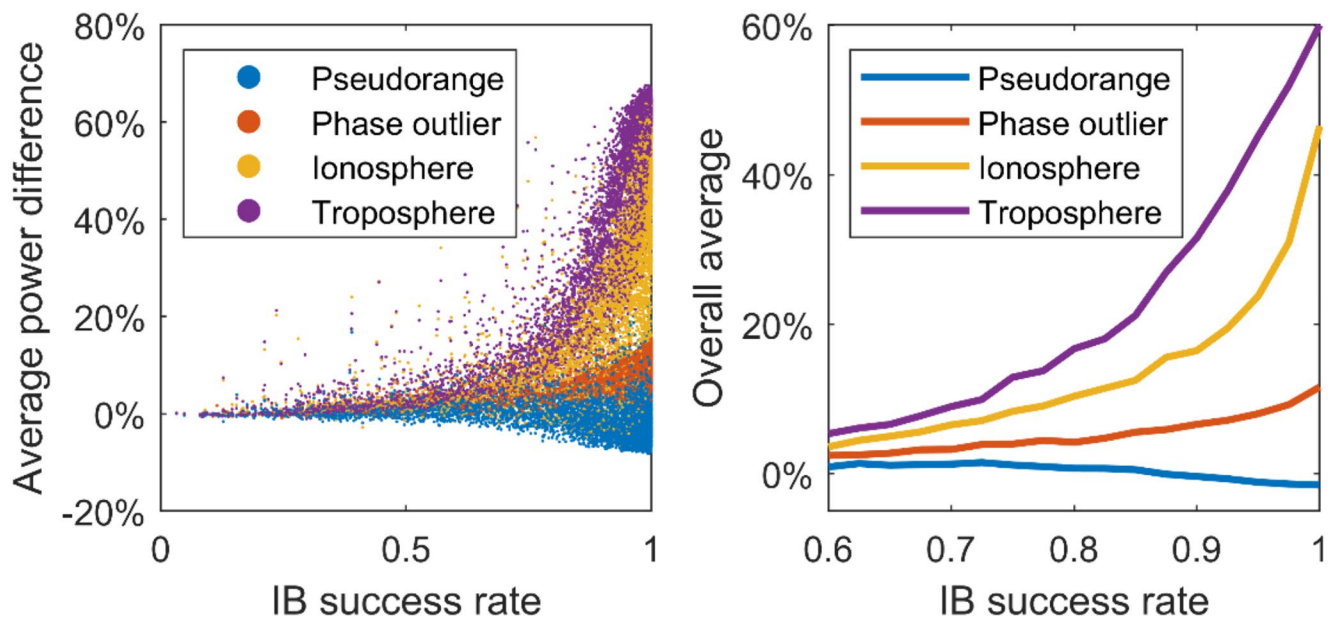
Fig. 11 Left graph shows average power difference vs. IB success rate of corresponding experiment for four misspecifications, which are computed from power functions with $\alpha = 0.01$. To obtain graph at right, we split success rate range $[0.6, 1]$ into 16 intervals and compute overall average of power difference within each interval

Table 4 Overall average of power difference in four success rate intervals

Misspecifications	$[0.775, 0.8]$	$[0.875, 0.9]$	$[0.925, 0.95]$	$[0.975, 1]$
Blunder in pseudorange	1%	-0%	-1%	-1%
Phase outlier	4%	7%	8%	12%
Ionosphere delay	10%	17%	24%	47%
Troposphere delay	17%	32%	45%	60%

Overall, the AR detector is better as it performs similarly to the AF detector for the blunder in pseudorange detection and delivers higher detection powers for phase outliers and atmosphere delays.

Multi-dimensional misspecifications

The performance of the AR detector for three multi-dimensional (qD, $q > 1$) misspecifications is evaluated and presented in this section. The simulation setup is the same as in the section 'Performance for dual-frequency model' where a dual-frequency (L1/L5) DD model is used. Detection for the following misspecifications is considered,

- carrier phase observables on L1 and L5 signals of satellite G05 are outliers (2D).
- blunders in pseudorange observables of two low-elevation satellites, G17 and G30 (4D).

- differential ionosphere delays affect the signals from the southwest direction (G05 and G24, 2D).

The power functions corresponding to the qD misspecifications are qD functions, which are evaluated on the qD grid points of $c = [\ c_1 \quad \cdots \quad c_q\ ]$. The power functions plotted in this section are evaluated with $\alpha = 0.01$ and $N_1 = N_2 = 2 \times 10^5$ for the Monte Carlo simulation of the critical values and power.

Figure 12 (left) shows the contour lines of the AF and AR detector for 2D phase outliers. The AR contour lines are more concentrated, indicating that the AR detector provides a higher detection power for phase outliers of the same size. Table 5 presents that the average power difference between the AR and AF detectors is 29% for $\alpha = 0.01$ in this experiment. The middle and right graphs in Fig. 12 show contour lines for two marginal distributions of the 4D blunder in pseudorange power function. Similar to the results in the one-dimensional case, the AR and AF detectors perform similarly. The average increase in the detection power is 6% for $\alpha = 0.01$, as is shown in Table 5. Figure 13 shows the power functions of the 2D differential ionosphere delay detection. We observe an oscillation of the AR power function, which is higher than that of the AF detector. Table 5 shows that the average improvement in the detection power is 68% for $\alpha = 0.01$. The AR detector provides higher detection powers than the AF detector, especially for ionosphere delay detection in multi-dimensional misspecification experiments.
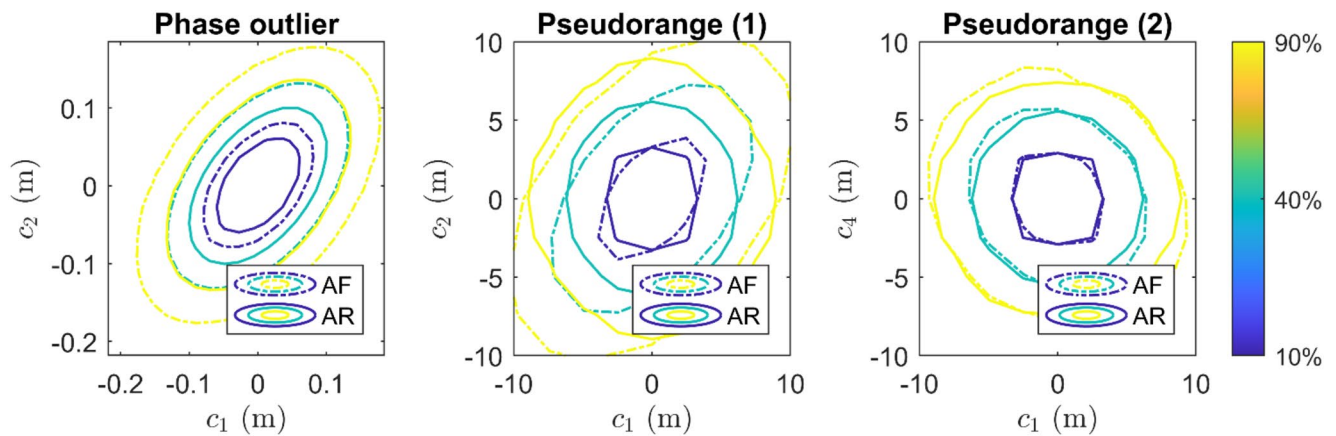
Fig. 12 Left graph shows contour lines of AF and AR power functions for 2D phase outlier detection, evaluated on 400 grid points. c_1 and c_2 regarding to phase outliers of G05 on L1 and L5. Middle and right graphs show contour lines for marginal distributions of the 4D blunder in pseudorange power function, which is evaluated on $13^4 = 28,561$ grid points. c_1 and c_2 regarding to blunder in pseudoranges of G17 on L1 and L5. c_4 is size of blunder in pseudorange of G30 on L5

Table 5 Average power difference for qD misspecification detection

Misspecification	$\alpha = 0.005$	$\alpha = 0.01$	$\alpha = 0.05$
Phase outlier	31%	29%	22%
Blunder in pseudorange	7%	6%	5%
Ionosphere delay	68%	68%	56%
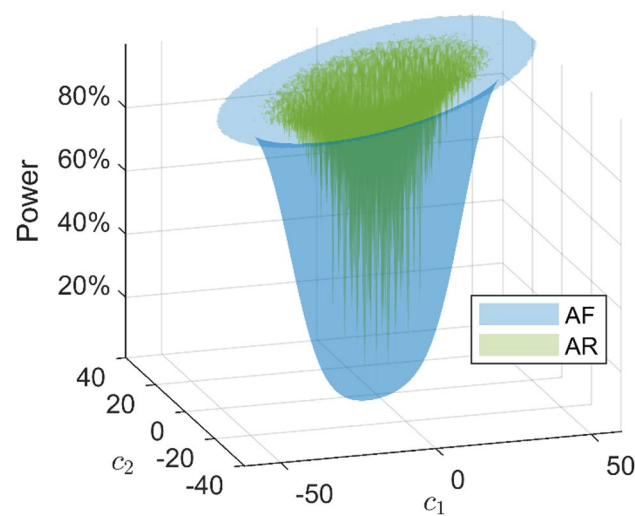
Fig. 13 AF and AR power functions for 2D ionosphere delay detection. Only powers smaller than 0.999 are shown for clarity of the figure. c_1 and c_2 are size (in TECU) of the differential ionosphere delays corresponding to G05 and G24

Summary and conclusions

In this contribution, we evaluated the performance of the ambiguity-resolved (AR) detector for the GNSS relative positioning model by obtaining the detection powers as functions of the misspecifications' size. We compared its performance with the ambiguity-float (AF) and ambiguity-known (AK) detectors.

We first presented the differential observation model under the null (H_0) and alternative (H_a) hypotheses. The model under H_a contains an additional term to account for misspecifications and is used to evaluate the detection power for the misspecification. The performance of detecting the blunders in pseudorange, phase outliers, ionosphere and troposphere delays, and constant bias was evaluated.

Then, we described how the detectors could be implemented and how to obtain power functions. The distribution of the AR test statistic cannot be evaluated analytically. We presented the procedure to obtain the AR critical value and detection power using Monte Carlo simulation. Moreover, we introduced how to assess the uncertainty of the simulated critical value and detection power.

We conducted simulation experiments to evaluate the performance of the AR detector. We started with three experiments of detecting one-dimensional misspecifications. We first conducted a single-frequency GPS experiment with an 80.5% double-differenced (DD) ambiguity resolution success rate, in which the distributions of the test statistics under H_0 and H_a were presented. It is shown by the power functions that the AR detector has larger detection power than the AF detector, even though the success rate is not close to one. Then, we conduct a dual-frequency GPS experiment with a 97.4% DD success rate. The difference between the AR and AF detection power of atmosphere delays is over 49% on average. In these two experiments, we also find that the AR power function differs from the AK power function. Hence, one should not predict the performance of the AR detector relying on the assumed probabilistic property of the AK detector as it is too optimistic.

Following that, we obtained power functions on 25 user locations, with five different observation models and 72 satellite geometries per location per model. For the experiments

with success rates larger than 97.5%, the AR detector performs similarly to the AF detector for the blunder in pseudorange detection, increases the detection power by 12% on average for the phase outlier detection, and increases the detection power remarkably by 46% and 60% on average for the ionosphere and troposphere detection, respectively. Finally, we presented experiments of detecting three multidimensional misspecifications. The AR detector performs better than the AF detector in all the experiments, especially for the two-dimensional ionosphere delay detection, where the powers are increased by 68% on average for $\alpha = 0.01$.

The simulation experiments show that the AR detector can provide a higher detection power than the AF detector even if the ambiguity resolution success rate is not close to one. For models with low success rates, although the ambiguity resolution cannot be used by the parameter estimation, the AR detector can contribute to model validation.

Author contributions CY performed the simulations and prepared the draft manuscript. CY, PT, and CT contributed to the discussion and interpretation of results. PT and CT supervised the research, reviewed and edited the manuscript.

Data availability The GNSS orbit product used in this research can be downloaded from https://cddis.nasa.gov/archive/gnss/products/.

Declarations

Competing interests The authors declare no competing interests.

References

Ahn YW, Kim D, Dare P (2006) Local tropospheric anomaly effects on GPS RTK performance. In: Proceedings of the 19th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2006), pp 1925–1935

Baarda W (1968) A testing procedure for use in geodetic networks. Netherlands Geodetic Commission, Publications on Geodesy, New Series 2, No.5, Delft, The Netherlands

Björck Å, Golub GH (1973) Numerical methods for computing angles between linear subspaces. Math Comput 27(123):579–594

Braasch MS (2017) Multipath. In: Teunissen PJG, Montenbruck O (eds) Springer handbook of global Navigation Satellite systems. Springer, Cham, pp 443–468

Chen W (2016) A remark on the GNSS single difference model with common clock scheme for attitude determination. J Appl Geodesy 10(3):167–173

Dao T et al (2024) Detecting outliers in local ionospheric model for GNSS precise positioning. GPS Solut 28(153)

Duan B, Hugentobler U, Montenbruck O (2024) A method to assess the quality of GNSS satellite phase bias products. GPS Solut 28(89). https://doi.org/10.1007/s10291-024-01634-6

El Masri M, Morio J, Simatos F (2021) Improvement of the cross-entropy method in high dimension for failure probability estimation through a one-dimensional projection without gradient estimation. Reliab Eng Syst Saf 216:107991

Euler HJ, Goad CC (1991) On optimal filtering of GPS dual frequency observations without using orbit information. Bull Geod 65:130–143

Feng S, Ochieng W, Moore T, Hill C, Hide C (2009) Carrier phase-based integrity monitoring for high-accuracy positioning. GPS Solut 13(1):13–22

Hernández-Pajares M, Juan JM, Sanz J, Aragón-Àngel À, García-Rigo A, Salazar D, Escudero M (2011) The ionosphere: effects, GPS modeling and the benefits for space geodetic techniques. J Geod 85(12):887–907

Hobinger T, Jakowski N (2017) Atmospheric signal propagation. In: Teunissen PJG, Montenbruck O (eds) Springer handbook of global Navigation Satellite systems. Springer, Cham, pp 165–194

Hofmann-Wellenhof B, Lichtenegger H, Collins J (2012) Global positioning system: theory and practice. Springer, New York

Huang J, van Graas F (2007) Comparison of tropospheric decorrelation errors in the presence of severe weather conditions in different areas and over different baseline lengths. J Navig 54(3):207–226. https://doi.org/10.1002/j.2161-4296.2007.tb00405.x

Jin XX, de Jong CD (1996) Relationship between satellite elevation and precision of GPS code observations. J Navig 49:253–265

Keong J (1999) GPS/GLONASS attitude determination with a common clock using a single difference approach. In: Proceedings of the 12th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 1999), pp 1941–1950

Kok JJ (1984) On data snooping and multiple outlier testing. US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Charting and Geodetic Services

Kouba J, Lahaye F, Tétreault P (2017) Precise point positioning. In: Teunissen PJG, Montenbruck O (eds) Springer Handbook of Global Navigation Satellite Systems, Chap. 25. Springer International Publishing, Cham, pp 723–752

Lawrence D, Langley RB, Kim D, Chan F (2006) Decorrelation of troposphere across short baselines. In: Proceedings of the IEEE/ION PLANS 2006, San Diego, CA, April, pp 94–102

Leick A, Rapoport L, Tatarnikov D (2015) GPS satellite surveying. Wiley, Hoboken

Metropolis N, Ulam S (1949) The Monte Carlo method. J Am Stat Assoc 44(247):335–341

Misra P, Enge P (2011) Global positioning system: signals, measurements, and performance. Ganga-Jamuna, Lincoln

Morio J, Balesdent M (2015) Estimation of rare event probabilities in complex aerospace and other systems: a practical approach.

Woodhead Publishing, Sawston. https://doi.org/10.1016/c2014-0-02344-1

Odijk D (2017) Positioning model. In: Teunissen PJG, Montenbruck O (eds) Springer Handbook of global Navigation Satellite systems. Springer, Cham, pp 605–638

Odijk D, Teunissen PJG (2008) ADOP in closed form for a hierarchy of multi-frequency single-baseline GNSS models. J Geod 82:473–492. https://doi.org/10.1007/s00190-007-0197-2

Odijk D, Zhang B, Khodabandeh A, Odolinski R, Teunissen P (2016) On the estimability of parameters in undifferenced, uncombined GNSS network and PPP-RTK user models by means of S-system theory. J Geodesy 90(1):15–44

Pullen S, Park YS, Enge P (2009) Impact and mitigation of ionospheric anomalies on ground-based augmentation of GNSS. Radio Sci 44(1):1–10

Serfling RJ (1980) Approximation theorems of mathematical statistics. Wiley, pp 74–103

Teunissen PJG (1995) The least squares ambiguity decorrelation adjustment: a method for fast GPS integer estimation. J Geod 70:65–82. https://doi.org/10.1007/bf00863419

Teunissen PJG (1997) A canonical theory for short GPS baselines. Part I: the baseline precision. J Geod 71(6):320–336

Teunissen PJG (1998) Success probability of integer GPS ambiguity rounding and bootstrapping. J Geod 72:606–612. https://doi.org/10.1007/s001900050199

Teunissen PJG (1999a) The probability distribution of the GPS baseline for a class of integer ambiguity estimators. J Geod 73(5):275–284. https://doi.org/10.1007/s001900050244

Teunissen PJG (1999b) An optimality property of the integer least-squares estimator. J Geod 73(11):587–593. https://doi.org/10.1007/s001900050269

Teunissen PJG (2000) Testing theory: an introduction. Series on mathematical geodesy and positioning. Delft University, Delft

Teunissen PJG (2002) The parameter distributions of the integer GPS model. J Geod 76:41–48. https://doi.org/10.1007/s001900100223

Teunissen PJG (2017) Carrier-phase integer ambiguity resolution. In: Teunissen PJG, Montenbruck O (eds) Springer handbook of global navigation satellite systems. Springer, Cham, pp 661–685

Teunissen PJG (2024) The ambiguity resolved detector: a new detector for the mixed-integer GNSS Model. J Geod 98(83):1–16

Teunissen PJG, Khodabandeh A (2015) Review and principles of PPP-RTK methods. J Geodesy 89(3):217–240. https://doi.org/10.1007/s00190-014-0771-3

Teunissen PJG, Montenbruck O (2017) Springer handbook of global navigation satellite systems. Springer, Cham

Tiberius CCJM (1998) Recursive data processing for kinematic GPS surveying. Publications on Geodesy, Netherlands Geodetic Commission

Verhagen AA (2003) On the approximation of the integer least-squares success rate: which lower or upper bound to use? J Glob Position Syst 2:117–124

Wang K, EI-Mowafy A, Qin W, Yang X (2022) Integrity monitoring of PPP-RTK positioning; part I: GNSS-based IM procedure. Rem Sens 14(1):44

Wanninger L (2004) Ionospheric disturbance indices for RTK and network RTK positioning. Paper presented at ION GPS/2004, Inst. of Navig., Long Beach, California

Yin C, Teunissen PJG, Tiberius CCJM (2024) Implementation of ambiguity-resolved detector for high-precision GNSS fault detection. In: Proceedings of the 37th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS + 2024), pp 2163–2174

Zhang W, Wang J, El-Mowafy A, Rizos C (2023) Integrity monitoring scheme for undifferenced and uncombined multi-frequency multi-constellation PPP-RTK. GPS Solut 27(2):68. https://doi.org/10.1007/s10291-022-01391-4