MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Bayesian Measurement Masks for GNSS Positioning

Greiff, Marcus; Di Cairano, Stefano; Berntorp, Karl

TR2024-172 December 18, 2024

Abstract

We propose a Bayesian measurement masking method for global navigation satellite system (GNSS) positioning to mitigate disturbances from multi-path biases and modeling errors. The method removes erroneous GNSS observations to improve performance in downstream positioning algorithms. The measurement masking is posed as a binary classification problem, and solved by sequentially determining the noise statistics of individual pseudo-range measurements in the GNSS observations. Bayesian probabilities of mismatching noise models inform when outlier events such as multipath or non-line-of-sight (NLOS) events occur. We report a classification F1-score of >0.99 when the modeling assumptions are satisfied, and >0.97 when realistic modeling errors are included, both for dynamic and static receiver motion models.

IEEE Conference on Decision and Control (CDC) 2024

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Bayesian Measurement Masks for GNSS Positioning

Marcus Greiff, Stefano Di Cairano, Karl Berntorp*

***Abstract*— We propose a Bayesian measurement masking method for global navigation satellite system (GNSS) positioning to mitigate disturbances from multi-path biases and modeling errors. The method removes erroneous GNSS observations to improve performance in downstream positioning algorithms. The measurement masking is posed as a binary classification problem, and solved by sequentially determining the noise statistics of individual pseudo-range measurements in the GNSS observations. Bayesian probabilities of mismatching noise models inform when outlier events such as multipath or non-line-of-sight (NLOS) events occur. We report a classification F1-score of >0.99 when the modeling assumptions are satisfied, and >0.97 when realistic modeling errors are included, both for dynamic and static receiver motion models.**

I. Introduction

High-precision Global Navigation Satellite System (GNSS) positioning is a critical component in modern infrastructure. However, in environments such as urban canyons where accuracy is most needed, high buildings and structures give rise to reflections of the GNSS signals, multi-path (MP) effects [1]. This leads to measurement biases that can be highly disruptive for positioning performance [2]. As a result, a vast body of literature has emerged on MP mitigation since its first description in the 1970s [3].

The receiver-internal MP mitigation is done at the lowest level of signal processing in the Delay-Lock Loops (DLLs) of the receiver, modifying the correlator used to compute the observables from the local pseudo-random noise (PRN) sequence. Some correlators include the high resolution [4] and vision correlators [5].

In terms of hardware mitigation, there have been developments in antenna design, such as the "choke ring", which is commonplace in modern GNSS equipment [6], and the less conventional dual- or multi-antenna array approaches [7], [8]. Such hardware modifications can further mitigate multi-path, but will similarly not result in perfect mitigation.

The third element to multi-path mitigation is *masking*. This includes studying signals from the GNSS chip set, such as carrier-to-noise density ratios (CN_0), signal-to-noise ratios (SNRs) and how they vary with satellite elevation [9]. A hard threshold is often implemented on these signals to mitigate multipath [10], a method used in high-end positioning libraries such as `RTKLlib` [11] and `PPPLib` [12]. Such heuristics emulate a binary classifier operating on a set of inputs (or features). Prior work proposes to learning this function by machine learning (ML) [13]. Our recent work [14] demonstrated that an auto-encoder with simple k-means clustering is well suited for the task, but requires domain adaptation to specific receivers and environments. When deploying receivers at scale, the thresholds in [10] or neural networks in [14] need to be adapted to specific devices (as the definition and quantization of CN_0 and SNR may differ between receivers), posing a significant challenge.

Contributions: We improve the masking by developing a method inspired by consistency checking [15], which does not rely on large amounts of training data as in typical ML [14]. We introduce the concept of a Bayesian measurement mask (BM), which optionally uses SNR and satellite elevation information, and detail two implementations using recursive estimation techniques: the interactive BM (IBM) based on interactive mixtures and the variational BM (VBM) using variational techniques. We compute a measurement mask based on *all measurements* and *prior knowledge* sequentially, instead of a classification based on features (such as satellite elevation and CN_0), or short sequences of measurements [10], [13], [14]. The resulting mask can be used as a noninvasive add-on to any GNSS positioning algorithm operating with raw GNSS measurements. Fig. 1 shows a schematic of the setup.

Notation: Vectors are denoted by $\boldsymbol{x} \in \mathbb{R}^N$ with $[\boldsymbol{x}]_i$ being the i^{th} element of $\boldsymbol{x}$. Matrices are indicated in bold, $\boldsymbol{X}$, and the element on row i and column j of $\boldsymbol{X}$ is $[\boldsymbol{X}]_{ij}$. The notation $\boldsymbol{I}_n$ means the $n \times n$ identity matrix and $\mathbf{1}_n$ is a vector of ones with dimension n. The notation $\underline{\boldsymbol{x}} \sim \mathrm{N}(\boldsymbol{x}|\boldsymbol{m}, \boldsymbol{P})$ indicates that $\underline{\boldsymbol{x}}$ is a Gaussian distributed random variable with mean $\boldsymbol{m}$ and covariance $\boldsymbol{P}$. The bar $\underline{(\cdot)}$ is used to indicate a random variable if necessary, but often dropped to clarify the presentation. Similarly, the notation $\underline{\boldsymbol{\Sigma}} \sim \mathrm{IW}(\boldsymbol{\Sigma}|\nu, \boldsymbol{V})$ indicates that $\underline{\boldsymbol{\Sigma}}$ is Inverse-Wishart distributed, with ν degrees of freedom and a scale matrix $\boldsymbol{V}$. Given the set of measurements $\boldsymbol{y}_{0:k} = \{\boldsymbol{y}_0, \dots, \boldsymbol{y}_k\}$, we let $p(\boldsymbol{x}_k|\boldsymbol{y}_{0:k})$ denote a marginal filtering posterior. Finally, we write the expectation of a random variable $\underline{\boldsymbol{x}} \sim p(\boldsymbol{x})$ as $\mathbb{E}_{\underline{\boldsymbol{x}}\sim p(\boldsymbol{x})}[\underline{\boldsymbol{x}}] = \int \boldsymbol{x} p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$ compactly as $\mathbb{E}_p[\boldsymbol{x}]$. In this notation, the Kullback-Leibler (KL) divergence of two densities $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$ is $\mathrm{KL}(p(\boldsymbol{x})||q(\boldsymbol{x})) = \mathbb{E}_p(\log(p(\boldsymbol{x})/q(\boldsymbol{x})))$.

II. Bayesian Measurement Masks

A mask is an operator that selects a subset of measurements to be used for filtering. For GNSS, this is a binary classifier that determines the presence or absence of modeling errors, true or false, respectively. With n satellites $\mathcal{S} = \{s \in \mathbb{N}_{[1,n]}\}$, we seek a function $f : \mathcal{S} \mapsto \bar{\mathcal{S}}$ where

$$\bar{\mathcal{S}} = \{s \in \mathcal{S} | s \text{ is not affected by multipath}\} \subseteq \mathcal{S}. \quad (2)$$

In conventional methods, f is memory-less and operates on single time-instances of certain signals, such as satellite

Mitsubishi Electric Research Labs (MERL), Cambridge, MA, USA.
*Karl Berntorp. Email: `karl.o.berntorp@ieee.org`.

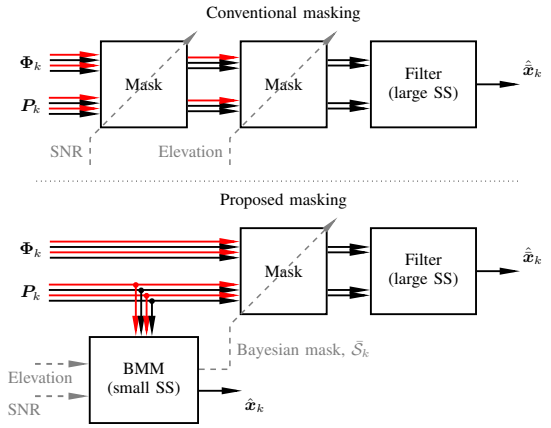

Fig. 1. Comparison between conventional masking and the proposed Bayesian masking. *Top:* In conventional masking, poor measurements (red) are removed by applying a sequence of masks based on CN_0 and satellite elevation angles with the goal of producing a better estimate in a filter operating with a relatively large state space (SS). *Bottom:* In the proposed method, a subset of the measurements are used along with CN_0 and satellite elevation angles to sequentially compute a measurement mask by posing a filtering problem with a smaller SS.

TABLE I

SUMMARY OF GNSS MODEL PARAMETERS.

Variable	Functional dep.	Description
$\rho^s_{\mathcal{R}}$	–	Euclidean distance from $\mathcal{R}$ to s
$dt_{\mathcal{R}}, dt^s$	–	Clock offset
$D_{\mathcal{R}}, D^s$	L_j and $\{P, \Phi\}$	Inter-frequency bias
N	L_j	Integer ambiguity bias
λ_j	L_j	$(= f_j^{-1})$ Carrier wavelength
f_j	L_j	$(= \lambda_j^{-1})$ Carrier frequency
$T^s_{\mathcal{R}}$	–	Tropospheric delay
$I^s_{\mathcal{R}}$	L_j	Ionospheric delay
$M^s_{\mathcal{R}}$	L_j and $\{P, \Phi\}$	Multipath effects
$\epsilon^s_{\mathcal{R}}$	L_j and $\{P, \Phi, \Gamma\}$	Uncorrelated Gaussian noise

elevation angles and CN_0 ratios [10], or small batches of measurement data [13]. The Bayesian masks instead at time step k compute the most likely mask to be true given the previous measurements and prior knowledge of, for instance, the dynamics of the receiver motion.

A. GNSS Measurements and Differencing

We consider standard GNSS measurements provided in the RINEX format [16], where $\mathcal{R}$ denotes a receiver, $\mathcal{B}$ denotes a base station, and the super-index $s \in \mathbb{N}_{>0}$ denotes a satellite. The measurements include a pseudo-range $P^s_{\mathcal{R}} \in \mathbb{R}$ computed by an auto-correlation on the pseudo-random code, the phase-range $\Phi^s_{\mathcal{R}} \in \mathbb{R}$ containing the integer ambiguity, and a Doppler measurement $\Gamma^s_{\mathcal{R}} \in \mathbb{R}$. The observation equations are given in (1), defined by signals summarized in Table I. Here, the functional dependency indicates channels on which the signal differs. For instance, the noise $\epsilon^{P,s}_{\mathcal{R}}$ between a satellite and receiver is realized differently on every combination of measurements $\{P, \Phi, \Gamma\}$ with a unique frequency band L_j. On the other hand, the initial oscillator phases only depend on the frequency band L_j, as they solely appear in the phase-range measurements.

To simplify, we let $\boldsymbol{p}_{\mathcal{R}}$, $\boldsymbol{p}_{\mathcal{B}}$, and $\boldsymbol{p}^s$ denote the positions of the receiver, base station, and satellites, respectively. Without loss of generality, we consider the L_1 frequency band. We let $\rho^s_{\mathcal{R}} = \|\boldsymbol{p}_{\mathcal{R}} - \boldsymbol{p}^s\|_2$ and collect the distances between a receiver $\mathcal{R}$ and the satellites in $\boldsymbol{\rho}_{\mathcal{R}} = (\rho^1_{\mathcal{R}}, ..., \rho^n_{\mathcal{R}})$, with analogous definitions for the observations $\boldsymbol{P}_{\mathcal{R}}$, $\boldsymbol{\Phi}_{\mathcal{R}}$, $\boldsymbol{\Gamma}_{\mathcal{R}}$ formed in $P^s_{\mathcal{R}}$, $\Phi^s_{\mathcal{R}}$, $\Gamma^s_{\mathcal{R}}$, and the noise as $\boldsymbol{\epsilon}^P_{\mathcal{R}}$, $\boldsymbol{\epsilon}^{\Gamma}_{\mathcal{R}}$, $\boldsymbol{\epsilon}^{\Phi}_{\mathcal{R}}$ formed in $\epsilon^{P,s}_{\mathcal{R}}$, $\epsilon^{\Phi,s}_{\mathcal{R}}$, $\epsilon^{\Gamma,s}_{\mathcal{R}}$. It is common to take differences of the observations to improve robustness to modeling errors [17], [18], and there are many ways of doing so. For simplicity, we define a single-difference operator

$$\text{SD}(\boldsymbol{P}) = \boldsymbol{P}_{\mathcal{R}} - \boldsymbol{P}_{\mathcal{B}}. \tag{3}$$

It is possible to take such differences without a base station

$$\overline{\text{SD}}(\boldsymbol{P}_{\mathcal{R}}) = \begin{bmatrix} \mathbf{1}_{n-1} & -\boldsymbol{I}_{n-1} \end{bmatrix} \boldsymbol{P}_{\mathcal{R}}, \tag{4}$$

or even take double differences (DD) $\overline{\text{SD}}(\text{SD}(\boldsymbol{P}))$. All of these operations are linear, and give rise to different correlation structures in the noise of the differenced observations.

B. The Large Filtering Problem and Conventional Masking

When posing filtering problems, it is common to assume uncorrelated Gaussian noise, compensate for the ionospheric and tropospheric delays using physical models, and take differences of the $\{\boldsymbol{P}_i, \boldsymbol{\Phi}_i, \boldsymbol{\Gamma}_i\}_{i \in \{\mathcal{R}, \mathcal{B}\}}$ signals in (1) to remove modeling errors. This results in a measurement model

$$p(\bar{\boldsymbol{y}}_k | \bar{\boldsymbol{x}}_k) = \text{N}(\bar{\boldsymbol{y}}_k | \bar{\boldsymbol{h}}(\bar{\boldsymbol{x}}_k), \bar{\boldsymbol{R}}_k), \tag{5}$$

where the state $\bar{\boldsymbol{x}}_k$ is high-dimensional (usually >100 states). The state contains the dynamic states of the receiver, including its position, as well as various real-valued biases. The state also contains SD or DD integer-valued ambiguity biases. This warrants special estimation algorithms, collectively referred to as *large filters* (LFs). Here, accuracy is contingent on fixating the integer ambiguities [17], [19]. The LFs typically include a subset of the measurements $\bar{\boldsymbol{y}}_k$, based on a set $\bar{\mathcal{S}} \subset \mathcal{S}$ of the satellites. Determining $\bar{\mathcal{S}}$ is a binary classification problem, commonly solved heuristically by masking out satellites based on elevation angles and CN_0 signals [10]. We let $\phi^s_k \in [0, \pi/2]$ denote the elevation angle of satellite s, and define the mask to include all measurements at a time step k formed with satellites satisfying $\phi^s_k > \xi$ for some threshold $\xi > 0$.

C. Bayesian Measurement Masking

If we are only concerned with establishing a measurement mask, then we can consider a *small filter* (SF) operating on a subset of the measurement information. For the purposes of this paper, the exact definitions of the LF is not important, and we will instead derive a smaller measurement model $p(\boldsymbol{y}_k | \boldsymbol{h}(\boldsymbol{x}_k), \boldsymbol{R}_k)$ and present SFs to solve the binary classification problem. To this end, we form single difference measurements where the terms related to I, T, D, dt in Table I are compensated for by deterministic models or estimates from the LF. Furthermore, we only consider the pseudo-range, as this obviates the need for estimating integer ambiguities. Consequently, we define a smaller model

$$\boldsymbol{y}_k = \text{SD}(\boldsymbol{\rho}_k) + \text{SD}(\boldsymbol{M}^P_k) + \text{SD}(\boldsymbol{\epsilon}^P_k). \tag{6}$$

$$P_i^s(L_j) = +\rho_i^s + c[dt_i - dt^s] + [D_i(L_j, P) - D^s(L_j, P)] + T_i^s + I_i^s(L_j) + M_i^s(L_j, P) + \epsilon_i^{P,s}(L_j) \quad (1a)$$

$$\Phi_i^s(L_j) = +\rho_i^s + c[dt_i - dt^s] + [D_i(L_j, \Phi) - D^s(L_j, \Phi)] + T_i^s - I_i^s(L_j) + M_i^s(L_j, \Phi) + \lambda_j N_i^s(L_j) + \epsilon_i^{\Phi,s}(L_j) \quad (1b)$$

$$\Gamma_i^s(L_j) = -\dot{\rho}_i^s - c[\dot{dt}_i - \dot{dt}^s] - \dot{T}_i^s + \dot{I}_i^s(L_j) - \dot{M}_i^s(L_j, \Phi) + \epsilon_i^{\Gamma,s}(L_j) \quad (1c)$$

The model (6) only includes the kinematic states of the receiver, $\boldsymbol{x}_k$, and any inconsistencies in the estimation model can be attributed to modeling errors or multi-path.

We envision a two-stage filtering solution, where the original LF remains intact but the conventional masks are replaced by a Bayesian estimator (see Fig. 1). The measurement mask using the posterior formed by the SF. We refer to this as Bayesian measurement masking (BM), and present two approaches in which it can be implemented. In Sec. III we propose an interactive mixture model filter [20] to find the most likely combination of nonzero elements $\{\boldsymbol{M}_k^P, \boldsymbol{M}_k^\Gamma\}$, forming the LF measurement mask for (5). This is referred to as the interactive Bayesian measurement mask (IBM). In Sec. IV, an alternative approach is taken, with the idea that modeling errors in (6) will increase the innovation errors for specific measurements. Here, a variational Bayes (VB) filter [21] is employed where we explicitly estimate the variance of the noise terms, and determine a mask for (5).

III. Bayesian Masks by Mixture Filtering

To implement a measurement mask using the interactive multiple modeling (IMM) framework, we make two assumptions, which reduce the computational complexity of the algorithm by bounding the number of models in the filter.

Assumption 1 *If multipath affects the code measurements from one satellite, it also affects the carrier-phase measurement from that same satellite. Formally,*

$$[\boldsymbol{M}_k^P]_i = 0 \Rightarrow [\boldsymbol{M}_k^\Phi]_i = 0, \quad (7a)$$

$$[\boldsymbol{M}_k^P]_i \neq 0 \Rightarrow [\boldsymbol{M}_k^\Phi]_i \neq 0. \quad (7b)$$

For non-line-of-sight (NLOS), Assumption 1 is obviously true. If we correctly detect that the code measurement is NLOS, then this is a property that depends on the world geometry and the relative location of the satellite, and the carrier phase would be affected. Thus, we do not need to consider the Doppler shifts, and can work with a small estimation model

$$p(\boldsymbol{y}_k|\boldsymbol{x}_k) = \mathrm{N}(\boldsymbol{y}_k|\boldsymbol{h}(\boldsymbol{x}_k), \boldsymbol{R}_k^m), \quad (8)$$

where $\boldsymbol{R}_k^m$ is the measurement noise covariance of the mth measurement model. To specify the set of measurement models and reduce its size, we make the following assumption.

Assumption 2 *A minimum of $\bar{n} \in \mathbb{N}_{[0,n]}$ satellites are free from multipath effects at any given time step k.*

We introduce Assumption 2 mainly as a means of controlling the computational scaling of the method. It can be set to $\bar{n} = 0$, thus covering all possible cases. Next, we consider the noise levels of the pseudorange measurements in (8). If we have compensated for the modeling errors correctly and there is no multipath, the noise associated with $\boldsymbol{y}_k$ is entirely determined by some nominal noise $\mathrm{N}(\boldsymbol{\epsilon}_k^P|\boldsymbol{0}, \boldsymbol{R}_k)$. Define

$$\mathcal{R}_k = \left\{ \boldsymbol{R}_k + R_H \mathrm{diag}(\boldsymbol{b}) \middle| \boldsymbol{b} \in \{0,1\}^n, \sum_{i=1}^n [\boldsymbol{b}]_i \geq \bar{n} \right\}, \quad (9)$$

for some R_H such that $R_H \boldsymbol{I} \succ \boldsymbol{R}_k$. In total, this amounts to $|\mathcal{R}_k| = \sum_{k=0}^{n-\bar{n}} \binom{n}{k}$ models in (8) characterized by $\boldsymbol{R}_k \in \mathcal{R}_k$, with each covariance matrix associated with a unique integer $m \in \{1, ..., |\mathcal{R}_k|\}$. We can then associate each model with a probability $w_k^m \in (0, 1)$, define a posterior

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{0:k}) = \sum_{m=1}^{|\mathcal{R}_k|} w_k \mathrm{N}(\boldsymbol{x}_k|\boldsymbol{m}_k^m, \boldsymbol{P}_k^m), \quad (10)$$

and use an interactive filter to update the mixture estimates $\{\boldsymbol{m}_k^m, \boldsymbol{P}_k^m\}_{m=1}^{|\mathcal{R}_k|}$ and the mode probabilities $\{w_k^m\}_{m=1}^{|\mathcal{R}_k|}$ sequentially from $p(\boldsymbol{x}_k|\boldsymbol{y}_{0:k-1})$ given a transition probability matrix (TPM) $\boldsymbol{\Pi} \in [0,1]^{|\mathcal{R}_k| \times |\mathcal{R}_k|}$ governing the Markov chain of the mode state. We find the most likely model as

$$m_k^\star = \arg\max_m w_k^m, \quad (11)$$

and include information from satellite s at a time step k in the LF. We define the masking function f at time step k by

$$\bar{\mathcal{S}}_k = \{s \in \mathcal{S} \mid [\boldsymbol{R}_k^{m_k^\star} - \boldsymbol{R}_k]_{ss} = 0\}. \quad (12)$$

A. Prior Information

As the IMM is derived in a Bayesian framework, it is possible to define priors both over the mode probabilities $\{w_0^m\}_{m=1}^{|\mathcal{R}_k|}$ and the individual densities $p(\boldsymbol{x}_0) = \mathrm{N}(\boldsymbol{x}_0|\boldsymbol{m}_0, \boldsymbol{P}_0)$. Similarly, we can encode assumed behaviors of the receiver motion through the model $p(\boldsymbol{x}_k|\boldsymbol{x}_{k+1})$. In this paper, we primarily consider a static receiver with a small random walk on its positions, and a dynamic receiver with a small random walk on its velocities.

Beyond standard modeling, it is also possible to encode behaviors about the multipath in the transition probabilities governing the Markov chain of the mode probabilities. For example, we can let the probability of transitioning from one mode to another based on the models in $\mathcal{R}_k$. To see this, let $\pi_{A\to B}^s$ denote the probability that a satellite s in a state $A \in \{0,1\}$ transitions to a state $B \in \{0,1\}$, with

$$\sum_{B \in \{0,1\}} \pi_{A\to B}^s = 1 \ \ \forall\ s = 1, .., n, \ \ A \in \{0,1\}. \quad (13)$$

Furthermore, assume that each satellite transitions to a state independent of any other satellite. Let $(i, j) \in \mathbb{N}^2_{[1,|\mathcal{R}_k|]}$ be two modes, and consider their models as specified with the associated binary variables $\boldsymbol{b}_i, \boldsymbol{b}_j$ in (9). We can then form

$$[\boldsymbol{\Pi}]_{ij} = \prod_{s=1}^n \pi_{[\boldsymbol{b}_i]_s \to [\boldsymbol{b}_j]_s}^s, \quad (14)$$

and encode relative probabilities of remaining in a multipath-free state directly in the masking method. It is also possible to incorporate other information, such as satellite elevation and CN_0 variation in the definition of $\pi^s_{A\to B}$. Algorithm (1) summarizes the approach.

B. Properties

If N denotes the dimension of $\boldsymbol{x}_k$, the time-complexity of the algorithm per time step is $O(N) = N^3 \sum_{k=0}^{n-\bar{n}} \binom{n}{k}$. For context, with a constant-velocity model in $\mathbb{R}^3$ we have $N = 6$, and the evaluation of a mask with $n = 13$ satellites and $\bar{n} = 5$ is then approximately equal to running a regular Kalman filter in the LF with 100 states (not including integer fixation). This computational scaling in n motivates a study of alternative methods of computing the measurement mask. For this reason, we resort to variational inference, and propose a second algorithm in Sec. IV.

Algorithm 1 Interactive Bayesian Masking (IBM).

Define prior $\boldsymbol{m}_0, \boldsymbol{P}_0$, TPM $\boldsymbol{\Pi}$, dimensions n, bound $\bar{n}$, nominal noise $\{\boldsymbol{R}_k\}_{k=1}^K$, and noise parameter R_H.
for $k = 1, ..., K$ **do**
  Receive: $t_k, \boldsymbol{y}_k$
  // Mixing and moment matching
  for $m = 1, ..., |\mathcal{R}_k|$ **do**
    $\bar{w}^m_{k-1} = \sum_{j=1}^{|\mathcal{R}_k|} [\boldsymbol{\Pi}]_{mj} \cdot w^j_{k-1}$
    $\bar{\boldsymbol{m}}^m_{k-1} = \sum_{j=1}^{|\mathcal{R}_k|} [\boldsymbol{\Pi}]_{mj} \cdot \frac{w^j_{k-1}}{\bar{w}^m_{k-1}} \boldsymbol{m}^j_{k-1}$
    $\bar{\boldsymbol{P}}^m_{k-1} = \sum_{j=1}^{|\mathcal{R}_k|} [\boldsymbol{\Pi}]_{mj} \cdot \frac{w^j_{k-1}}{\bar{w}^m_{k-1}} [\boldsymbol{P}^j_{k-1} - (\bar{\boldsymbol{m}}^j_{k-1} - \bar{\boldsymbol{m}}^m_{k-1})(\bar{\boldsymbol{m}}^j_{k-1} - \bar{\boldsymbol{m}}^m_{k-1})^\top]$
  end for
  Set $h_k = t_k - t_{k-1}$
  for $m = 1, ..., |\mathcal{R}_k|$ **do**
    // Time prediction
    Predict $\{\boldsymbol{m}^m_{k|k-1}, \boldsymbol{P}^m_{k|k-1}\}$ from $\{\bar{\boldsymbol{m}}^m_{k-1}, \bar{\boldsymbol{P}}^m_{k-1}\}$
    // Measurement update
    Compute $\boldsymbol{H}^m_k = (\partial/\partial \boldsymbol{x}_k)\boldsymbol{h}(\boldsymbol{x}_k)|_{\boldsymbol{x}_k = \boldsymbol{m}^m_{k|k-1}}$
    $\boldsymbol{E}^m_k = \boldsymbol{y}_k - \boldsymbol{h}(\boldsymbol{m}^m_{k|k-1})$
    $\boldsymbol{S}^m_k = \boldsymbol{H}^m_k \boldsymbol{P}^m_{k|k-1} (\boldsymbol{H}^m_k)^\top + \boldsymbol{R}^m_k$
    $\boldsymbol{K}^m_k = \boldsymbol{P}^m_{k|k-1} (\boldsymbol{H}^m_k)^\top (\boldsymbol{S}^m_k)^{-1}$
    $\boldsymbol{m}^m_k = \boldsymbol{m}_{k|k-1} + \boldsymbol{K}^m_k \boldsymbol{E}^m_k$
    $\boldsymbol{P}^m_k = (\boldsymbol{I} - \boldsymbol{K}^m_k (\boldsymbol{H}^m_k)^\top) \boldsymbol{P}^m_{k|k-1}$
    $w^m_k = \mathrm{N}(\boldsymbol{y}_k | \boldsymbol{h}(\boldsymbol{m}^m_{k|k-1}), \boldsymbol{S}^m_k) \cdot \bar{w}^m_{k|k-1}$
  end for
  // Compute mask (satellites to include)
  $m^\star = \arg\max_m w^m_k$
  $\bar{\mathcal{S}}_k = \{s \in \mathcal{S} \mid [\boldsymbol{R}^{m^\star}_k - \boldsymbol{R}_k]_{ss} = 0\}$
end for

IV. Bayesian Masks by Variational Inference

Another popular method for estimating noise statistics in Gaussian state-space models is by the VB methods [21]. For Gaussian densities, we define an IW-distribution over the noise covariance matrices, which is a conjugate prior to the covariance matrix of a multivariate normal distribution. However, the sum of IW random variables are not IW distributed. Hence, to make the methods in [21] applicable to determining noise of single satellites under SD operations, we opt to estimate the SD-covariance directly. Specifically, we consider a measurement model

$$p(\boldsymbol{y}_k | \boldsymbol{x}_k, \boldsymbol{\Sigma}_k) = \mathrm{N}(\boldsymbol{y}_k | \boldsymbol{h}(\boldsymbol{x}_k), \boldsymbol{\Sigma}_k). \tag{15}$$

In VB, the idea is to approximate the posterior with factors

$$p(\boldsymbol{x}_k, \boldsymbol{\Sigma}_k | \boldsymbol{y}_{0:k}) \approx \underbrace{\mathrm{N}(\boldsymbol{x}_k | \boldsymbol{m}_k, \boldsymbol{P}_k)}_{\triangleq q_{\boldsymbol{x}}(\boldsymbol{x}_k)} \underbrace{\mathrm{IW}(\Sigma_k | \nu_k, \boldsymbol{V}_k)}_{\triangleq q_{\boldsymbol{\Sigma}}(\boldsymbol{\Sigma}_k)}, \tag{16}$$

and solve a variational optimization problem

$$\min_{q_{\boldsymbol{x}}, q_{\boldsymbol{\Sigma}}} \mathrm{KL}(q_{\boldsymbol{x}}(\boldsymbol{x}_k) q_{\boldsymbol{\Sigma}}(\boldsymbol{\Sigma}_k) || p(\boldsymbol{x}_k, \boldsymbol{\Sigma}_k | \boldsymbol{y}_{0:k})). \tag{17}$$

By variational calculus [22], the extremals are

$$q_{\boldsymbol{x}}(\boldsymbol{x}_k) \propto \exp(\mathbb{E}_{\boldsymbol{q}_{\boldsymbol{\Sigma}}}[p(\boldsymbol{x}_k, \boldsymbol{\Sigma}_k, \boldsymbol{y}_k | \boldsymbol{y}_{0:k-1})]), \tag{18a}$$

$$q_{\boldsymbol{\Sigma}}(\boldsymbol{\Sigma}_k) \propto \exp(\mathbb{E}_{\boldsymbol{q}_{\boldsymbol{x}}}[p(\boldsymbol{x}_k, \boldsymbol{\Sigma}_k, \boldsymbol{y}_k | \boldsymbol{y}_{0:k-1})]). \tag{18b}$$

Eq. (18) can be evaluated to a local maxima by fixed-point iterations, as the expectations in (18) are known in closed form given the assumed form of the posterior in (16) [21].

A. Motion models

To implement the VBM, we require motion models for the Gaussian states and the IW-parameters. For the Gaussian distribution of the receiver states, we consider the model

$$p(\boldsymbol{x}_k | \boldsymbol{x}_{k-1}) = \mathrm{N}(\boldsymbol{x}_k | \boldsymbol{A}_k \boldsymbol{x}_{k-1}, \boldsymbol{Q}_k), \tag{19}$$

with the same static and dynamic model as in the IBM. For the noise-covariance dynamics, we define a motion model

$$p(\boldsymbol{\Sigma}_k | \boldsymbol{\Sigma}_{k-1}) = \mathrm{IW}(a_k \nu_{k-1} + b_k(n+1), a_k \boldsymbol{V}_{k-1}), \tag{20a}$$

$$a_k = \exp(-h\tau^{-1}), \tag{20b}$$

$$b_k = \exp(-h\tau^{-1})(\exp(h\tau^{-1}) - 1), \tag{20c}$$

with a parameter $\tau > 0$ that can be interpreted as a time-constant. Eq. (20) is obtained by interpreting the IW-parameters as evolving according to an ordinary differential equation that is discretized at a sampling period of h. This effectively implements an exponential forgetting factor, which increases as $\tau \to 0$. In the limit $\tau \to \infty$, the IW-parameters are constant in the prediction, and unlike the model in [21], the model (20) is independent on h and can be implemented for variable rates.

B. Computing the measurement mask

Given the linear measurement model and assumptions of independence of the code (pseudorange) measurements, the noise-covariance metric should have a structure $\boldsymbol{\Sigma}_+ = \sigma^2_+ \mathbf{1}\mathbf{1}^\top$, being the nominal noise from the positive satellite in the SD scheme, and $\bar{\boldsymbol{\Sigma}}_k = \mathrm{diag}(\sigma^2_{2,k}, ..., \sigma^2_{n,k})$, being the contribution from each individual satellite (compare (8) and (15)). To assess the satellites that are afflicted by multipath, we can implement a threshold on the quantiles of the IW posterior, specifically over the diagonals of the measurement-noise covariance matrix. Specifically,

$$\mathbb{E}_{q_{\boldsymbol{\Sigma}}}[\boldsymbol{\Sigma}_k] = (\nu_k - n - 1)^{-1} \boldsymbol{V}_k, \tag{21}$$

and we determine the mask by a threshold $\eta > 0$ on (21). That is, we define the masking function f at time step k by

$$\bar{\mathcal{S}}_k = \{s \in \mathcal{S} \mid [(\nu_k - n - 1)^{-1}\boldsymbol{V}_k]_{ss} \leq \eta\}. \tag{22}$$

C. Prior information

As in the IBM, we can include priors in the form of initial estimates and initial measurement-noise covariances. We can include behaviors of the multipath effects by the parameter τ, which should be chosen small if the multipath biases vary rapidly. However, this parameter must be chosen in tandem with the masking threshold η. It is less obvious how to include other signals, such as satellite elevation and CN_0 signals. One possibility is by modifying (20) to drive the measurement-noise covariance to some target noise, and model this target noise based on assumptions on how it varies with elevation angle and CN_0. Algorithm 2 summarizes the proposed method.

D. Properties

The appeal of the variational framework for BM is that we update a single model through a set of fixed-point iterations, instead of updating a large number of models in the IBM. When capping the number of iterations at $I_{\max}$, the VBM is approximately a factor $(\sum_{k=0}^{n-\bar{n}} \binom{n}{k}) I_{\max}^{-1}$ faster than the IBM, thus making it a suitable option for applications involving a large number of satellites from multiple constellation and multiple frequency bands, where the number of code measurements is relatively large. However, since interactive mixture filters tend to outperform their variational filtering counterparts, we expect performance in the classification to be slightly worse in the VBM. Similar to the IBM, it is possible to determine certainty of the mask from the IW-posterior. One option is to relate the mask $\bar{\mathcal{S}}_k$ to a nominal covariance matrix in $\boldsymbol{R}_k^m \in \mathcal{R}_k$, and compute a relative likelihood against the mode of the IW-distribution, which is $(\nu_k + n + 1)^{-1}\boldsymbol{V}_k$. That is, we can compute a ratio

$$C_k = \frac{\mathrm{IW}(\boldsymbol{R}_k^m | \nu_k, \boldsymbol{V}_k)}{\mathrm{IW}((\nu_k + n + 1)^{-1}\boldsymbol{V}_k | \nu_k, \boldsymbol{V}_k)} \in [0, 1], \tag{23}$$

and use (23) to indicate the quality of the computed measurement mask. This information can also be leveraged in the LF, which typically has internal logic (resets, integer fixation) that depends on satellite health.

Algorithm 2 Variational Bayesian Masking (VBM).

Define prior $\boldsymbol{m}_0, \boldsymbol{P}_0, \nu_0^-, \boldsymbol{V}_0^-$, time-constant τ, dimensions n, maximum iterations $I_{\max}$, and threshold η.
for $k = 1, ..., K$ **do**
  Receive: $t_k, \boldsymbol{y}_k$
  // Time prediction
  Set $h_k = t_k - t_{k-1}$
  $\{\boldsymbol{m}_k^-, \boldsymbol{P}_k^-\} \leftarrow \mathrm{predict}(\{\boldsymbol{m}_{k-1}, \boldsymbol{P}_{k-1}\}, h_k)$ by (19)
  $\{\nu_k^-, \boldsymbol{V}_k^-\} \leftarrow \mathrm{predict}(\{\nu_{k-1}, \boldsymbol{V}_{k-1}\}, h_k)$ by (20)
  // Measurement update
  Compute $\boldsymbol{H}_k = (\partial/\partial \boldsymbol{x}_k)\boldsymbol{h}(\boldsymbol{x}_k)|_{\boldsymbol{x}_k = \boldsymbol{m}_k^-}$
  Set $\{\boldsymbol{m}_k^{(0)}, \boldsymbol{P}_k^{(0)}\} \leftarrow \{\boldsymbol{m}_k^-, \boldsymbol{P}_k^-\}$
  Set $\{\nu_k, \boldsymbol{V}_k^{(0)}\} \leftarrow \{\nu_k^- + 1, \boldsymbol{V}_k^-\}$
  Set $j = 0$
  while not converged and $j < I_{\max}$ **do**
    $\hat{\boldsymbol{\Sigma}}_k^{(j+1)} = \frac{1}{\nu_k - n - 1}\boldsymbol{V}_k^{(j)}$
    $\boldsymbol{S}_k^{(j+1)} = \boldsymbol{H}_k \boldsymbol{P}_k^- \boldsymbol{H}_k^\top + \hat{\boldsymbol{\Sigma}}_k^{(j+1)}$
    $\boldsymbol{K}_k^{(j+1)} = \boldsymbol{P}_k^- \boldsymbol{H}_k^\top (\boldsymbol{S}_k^{(j+1)})^{-1}$
    $\boldsymbol{m}_k^{(j+1)} = \boldsymbol{m}_k^- + \boldsymbol{K}_k^{(j+1)}(\boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{m}_k^-)$
    $\boldsymbol{P}_k^{(j+1)} = \boldsymbol{P}_k^- - \boldsymbol{K}_k^{(j+1)} \boldsymbol{S}_k^{(j+1)} (\boldsymbol{K}_k^{(j+1)})^\top$
    $\boldsymbol{E}_k^{(j+1)} = (\boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{m}_k^{(j)})(\boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{m}_k^{(j)})^\top$
    $\boldsymbol{V}_k^{(j+1)} = \boldsymbol{V}_k^- + \boldsymbol{E}_k^{(j+1)} + \boldsymbol{H}_k \boldsymbol{P}_k^{(j)} \boldsymbol{H}_k^\top$
    $j = j + 1$
  end while
  // Update statistics
  $\{\boldsymbol{m}_k, \boldsymbol{P}_k, \boldsymbol{V}_k\} \leftarrow \{\boldsymbol{m}_k^{(j)}, \boldsymbol{P}_k^{(j)}, \boldsymbol{V}_k^{(j)}\}$
  // Compute mask (measurements to include)
  $\bar{\mathcal{S}}_k = \{s \in \mathbb{N}_{[1,n]} \mid [(\nu_k - n - 1)^{-1}\boldsymbol{V}_k]_{ss} \leq \eta\}$
end for

V. Numerical Results

We consider two simulation settings: an ideal setting and a nonideal setting. In the former, the multipath effects are modeled by an increasing variance in the measurement noise, with the data generated by selecting a model randomly from $\mathcal{R}_k$, and changing the model every 10s. In the nonideal setting, the outliers are modeled by selecting a model randomly from $\mathcal{R}_k$, but multiplying the outlier variance by a random factor $f_m \sim \mathcal{U}([1, 5.0])$ and including a bias sampled uniformly from $\boldsymbol{b}_m \sim \mathcal{U}([-10, 10]^n)$. In summary, the noises in the synthetic data for this study are sampled from

- $\mathrm{N}(\boldsymbol{\epsilon}_k^P | \boldsymbol{0}, \boldsymbol{R}_k^{m_\ell})$ (ideal case);
- $\mathrm{N}(\boldsymbol{\epsilon}_k^P | \boldsymbol{b}_{m_\ell}, \boldsymbol{R}_k + f(\boldsymbol{R}_k^{m_\ell} - \boldsymbol{R}_k))$ (nonideal case);

where $\{m_\ell\}_{\ell=1}^{\lceil hK/10 \rceil}$ is a sequence of measurement models, with a new model (and associated bias) chosen every 10s. Furthermore, we consider two motion models: a static motion model, with a random walk on the positions of the receiver, and a dynamic model, with a random walk on the velocities of the receiver. The motion models and parameters for the simulations are provided in the appendix.

A. Performance Metrics

We are primarily interested in binary classification performance, as the job of the BM is to provide a measurement mask to the LF (see Fig. 1). To study this, we consider individual satellites, and distinguish between an outlier event (true) and the absence of an outlier event (false). Based on this, we compute the precision, recall, and the F1-score (their harmonic mean) and use these metrics to assess the classification performance. However, it is also relevant to understand if the filters are working as intended in the context of mismatching measurement models (nonideal case). We report the BMM estimates in the squared error $\mathrm{SE}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k) = \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k\|_2^2$ and the covariance trace $\mathrm{CT}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k) = \mathrm{Trace}(\boldsymbol{P}_k)$. Finally, the BMs report classification confidence by construction. For example, in the

TABLE II

CLASSIFICATION PERFORMANCE WITH **STATIONARY RECEIVER**

Method	Case(s)	Precision (↑)	Recall (↑)	F1-score (↑)
Elev. mask	Both	0.745	0.585	0.655
IBM	Ideal	0.990	**0.989**	**0.990**
IBM	Nonideal	0.987	0.941	0.963
VBM	Ideal	**0.991**	0.964	0.977
VBM	Nonideal	0.969	0.980	0.974

TABLE III

CLASSIFICATION PERFORMANCE WITH **MOVING RECEIVER**

Method	Case(s)	Precision (↑)	Recall (↑)	F1-score (↑)
Elev. mask	Both	0.745	0.585	0.655
IBM	Ideal	0.990	**0.990**	**0.990**
IBM	Nonideal	0.990	0.977	0.984
VBM	Ideal	0.976	0.986	0.981
VBM	Nonideal	**0.992**	0.875	0.930

IBM, we can measure the likelihood of the mask as

$$w_k^{\mathrm{ML}} = \max_{m \in |\mathcal{R}_k|} w_k^m \in [0, 1]. \tag{24}$$

B. Quantitative Study of Bayesian Masks

To assess the performance of the Bayesian methods, we perform a quantitative study where the Bayesian masks are compared with heuristics. We consider both the ideal and nonideal settings, the IBM and VBM, and the non-Bayesian method that applies a measurement mask based on the satellite elevation angle described in Sec. II-B. In this study, we consider a binary classification problem, where a satellite in a state 0 (false) is indicative of no multipath, and a state of 1 (true) is indicative of multipath. Here, false positives captured by the precision measure is of significant importance, indicating that an outlier-corrupted measurement is passed through the mask. To study this comprehensively, we compute the precision, recall, and F1 score from 100 Monte-Carlo runs, see Tables II and III, for the static and dynamic positioning case, respectively.

We note impressive prediction accuracy with the IBM method over all tested cases, performing en par with sophisticated learning-based methods [14] but without the need for domain adaptation to specific environments or receivers. In the elevation mask, performance of the classification solely depends on the satellite positions, and the method is therefore agnostic to any biases added to the measurements. This is also true for the masks in `RTKLlib` [11] and `PPPLib` [12].

VI. CONCLUSIONS

We present a method for multipath mitigation in GNSS positioning using Bayesian estimation techniques, and assume that multipath on the code measurements indicate multipath on the carrier-phase and Doppler measurements as well. Based on this, we formulate a smaller estimation problem, and propose two Bayesian methods to perform masking.

The VBM performs almost as well as the IBM, but yields significant improvements in computational scaling, making it more suitable for large, multi-constellation, and multi-band GNSS positioning problems. Importantly, the proposed masks depend on *all* code measurements, and not on a short history of measurements or auxiliary indicators such as elevation angles and CN_0 density ratios. In the proposed methods, assumed knowledge of these indicators can be incorporated as priors.

REFERENCES

[1] F. Fuschini, H. El-Sallabi, V. Degli-Esposti, L. Vuokko, D. Guiducci, and P. Vainikainen, "Analysis of multipath propagation in urban environment through multidimensional measurements and advanced ray tracing simulation," *IEEE Trans. Antennas and Propagation*, vol. 56, no. 3, pp. 848–857, 2008.

[2] M. Smyrnaios, S. Schn, M. Liso, and S. Jin, "Multipath propagation, characterization and modeling in GNSS," *Geodetic sciences-observations, modeling and applications*, pp. 99–125, 2013.

[3] L. Hagerman, "Effects of multipath on coherent and noncoherent PRN ranging receiver," *Aerosp. TOR-0073 (3020-03)-3*, 1973.

[4] G. A. McGraw and M. S. Braasch, "GNSS multipath mitigation using gated and high resolution correlator concepts," in *Proc. National techn. meeting of the institute of navigation*, 1999, pp. 333–342.

[5] P. C. Fenton and J. Jones, "The theory and performance of NovAtel Inc.'s vision correlator," in *Proc. Int. Techn. Meeting Satellite Division of The Institute of Navigation*, 2005, pp. 2178–2186.

[6] T. L. Blakney, D. D. Connell, B. J. Lamberty, and J. R. Lee, "Broadband antenna structure having frequency-independent, low-loss ground plane," Aug. 26 1986, uS Patent 4,608,572.

[7] P. D. Groves, Z. Jiang, B. Skelton, P. A. Cross, L. Lau, Y. Adane, and I. Kale, "Novel multipath mitigation methods using a dual-polarization antenna," in *Proc. Int. Techn. mtng. ION*, 2010, pp. 140–151.

[8] J. K. Ray, M. E. Cannon, and P. Fenton, "GPS code and carrier multipath mitigation using a multiantenna system," vol. 37, no. 1, pp. 183–195, 2001.

[9] A. Bilich, K. M. Larson, and P. Axelrad, "Modeling GPS phase multipath with SNR: Case study from the salar de uyuni, boliva," *J. Geophysical Research: Solid Earth*, vol. 113, no. B4, 2008.

[10] H. Tokura and N. Kubo, "Efficient satellite selection method for instantaneous RTK-GNSS in challenging environments," *Trans. Japan Soc. for Aero. Space Sciences*, vol. 60, no. 4, pp. 221–229, 2017.

[11] T. Takasu and A. Yasuda, "Development of the low-cost rtk-gps receiver with an open source program package rtklib," in *International symposium on GPS/GNSS*, vol. 1. International Convention Center Jeju Korea Seogwipo-si, Republic of Korea, 2009, pp. 1–6.

[12] C. Chen and G. Chang, "PPPLib: an open-source software for precise point positioning using GPS, BeiDou, Galileo, GLONASS, and QZSS with multi-frequency observations," GPS Solutions, 2021, Art. No.18.

[13] C. Savas and F. Dovis, "Multipath detection based on K-means clustering," in *Proc. Int. Techn. Meeting Satellite Division of The Institute of Navigation*, 2019, pp. 3801–3811.

[14] R. Zawislak, M. Greiff, K. J. Kim, K. Berntorp, S. Di Cairano, M. Konishi, K. Parsons, P. V. Orlik, and Y. Sato, "GNSS multipath detection aided by unsupervised domain adaptation," in *Proc. Int. Techn. mtng. ION*, 2022, pp. 2127–2137.

[15] G. Zhang, W. Wen, and L.-T. Hsu, "A novel GNSS based v2v cooperative localization to exclude multipath effect using consistency checks," in *Position, Location and Nav. Symp.*, 2018, pp. 1465–1472.

[16] RINEX Specification, "RINEX: The Receiver Independent Exchange Format, Version 3.03," International GNSS Service, Tech. Rep., 2015. [Online]. Available: https://files.igs.org/pub/data/format/rinex303.pdf

[17] K. Berntorp, A. Weiss, and S. Di Cairano, "Integer ambiguity resolution by mixture Kalman filter for improved GNSS precision," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 4, pp. 3170–3181, 2020.

[18] M. Greiff, S. Di Cairano, K. J. Kim, and K. Berntorp, "A system-level cooperative multiagent GNSS positioning solution," *IEEE Trans. Control Syst. Technol.*, vol. 32, no. 1, pp. 158–173, 2024.

[19] P. J. Teunissen, "The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation," *J. Geodesy*, vol. 70, no. 1, pp. 65–82, 1995.

[20] H. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Autom. Control*, 1988.

[21] S. Särkkä and J. Hartikainen, "Non-linear noise adaptive kalman filtering via variational Bayes," in *Int. Workshop Machine Learning for Signal Process.*, 2013, pp. 1–6.

[22] T. S. Jaakkola, "Variational methods for inference and estimation in graphical models," Ph.D. dissertation, MIT, 1997.