

Convex Reformulation of Information Constrained Linear State Estimation with Mixed-Binary Variables for Outlier Accommodation

Wang Hu, Zeyi Jiang, Hamed Mohsenian-Rad, and Jay A. Farrell

Abstract—This article considers the challenge of accommodating outlier measurements in state estimation. The Risk-Averse Performance-Specified (RAPS) state estimation approach addresses outliers as a measurement selection Bayesian risk minimization problem subject to an information accuracy constraint, which is a non-convex optimization problem. Prior explorations into RAPS rely on exhaustive search, which becomes computationally infeasible as the number of measurements increases. This paper derives a convex formulation for the RAPS optimization problems via transforming the mixed-binary variables into linear constraints. The convex reformulation herein can be solved by convex programming toolboxes, significantly enhancing computational efficiency. We explore two specifications: Full-RAPS, utilizing the full information matrix, and Diag-RAPS, focusing on diagonal elements only. The simulation comparison demonstrates that Diag-RAPS is faster and more efficient than Full-RAPS. In comparison with Kalman Filter (KF) and Threshold Decisions (TD), Diag-RAPS consistently achieves the lowest risk, while achieving the performance specification when it is feasible.

I. INTRODUCTION

The rapid decrease in the cost and size of sensors has led to an abundant number of measurements, far exceeding the number required for state observability. For instance, position determination using Global Navigation Satellite Systems (GNSS) is possible with pseudorange measurements from four spatially diverse satellites [1]. There are now four globally operational satellite constellations (e.g., BeiDou, GPS, Galileo, GLONASS) each providing pseudorange measurements on multi-frequencies. The number of GNSS measurements per epoch, under open-sky conditions, is predicted to increase toward 100 in the near future [2], [3]. Similarly, feature-based navigation only requires four features with diverse directions for observability [4]–[7], while cameras can provide hundreds of features per image [8]. Such applications, where the number of available measurements far exceeds the number required for observability are *signal-rich*.

For signal-rich applications, there is an important performance versus Bayesian risk trade-off. If all measurements were outlier-free with known accuracy, then using all measurements within a maximum a posteriori framework would yield the optimal state estimation performance [9], [10]. However, in typical applications some measurements are affected by outliers. Therefore, each measurement that is used has the potential to provide enhanced estimation

accuracy as quantified through the state error covariance matrix, if the measurement is outlier-free. Alternatively, there is the risk that outlier affected measurements will corrupt the state estimate while causing the estimator to be overly confident in that corrupted estimate (i.e., the error covariance matrix is too small).

This trade-off and methods to directly address it are critically important because at any measurement epoch several measurements may contain outliers. When the number of measurements that are used is large, adding one additional valid measurement increases the accuracy only marginally (i.e., covariance decreases proportional to $\frac{1}{m}$, where m is the number of measurements used); however, use of a single outlier-affected measurement can destroy the validity of all subsequent information processing by corrupting both the mean and covariance of the state estimate.

State estimation using discrete-time measurements involves two steps. The time update propagates the posterior state estimate from the previous measurement epoch across the time span between measurements, to provide a prior state estimate at the next measurement epoch. The measurement update incorporates the measurement information to correct the prior state estimate to produce a posterior state estimate at the measurement time. Because outliers only affect the measurement update, this paper focuses entirely on the problem of measurement selection for the measurement update, with a focus on signal-rich environments.

Outliers are the measurements that are inconsistent with the model and the remainder of the set of measurements [11]. Accommodation of outliers in state estimation is a well-studied problem. Standard methods for state estimation and control (see e.g.: [12], [13]) evaluate measurement residuals against their expected values using a fixed or an adaptive threshold test, as will be reviewed in Section III-A. More advanced outlier accommodation methods have been presented in other applications [14]–[17], such as Least Squares-based linear model fitting. These methods either detect and ignore outliers or de-weight measurements with large residuals. None of these approaches consider the risk-versus-performance trade-off involved with all measurements.

RAPS is an alternative approach to addressing outliers in state estimation [18]–[20]. This approach, reviewed in Section III-B, selects measurements to minimize a cost function that quantifies Bayesian risk while satisfying a performance constraint specified through the information matrix. Prior work on RAPS presented the optimization problem in its natural non-convex form. To achieve the global optimal solution, an Depth-First Search (DFS) approach was suggested

W. Hu, Z. Jiang, H. Mohsenian-Rad and J. A. Farrell are with the Department of Electrical and Computer Engineering, University of California, Riverside, CA 92521, USA. (e-mail: {whu027, zjian037}@ucr.edu, {hamed, farrell}@ece.ucr.edu).

to examine all 2^m combinations of the binary measurement selection vector [21]. This strategy becomes computationally infeasible as m increases. This article presents and demonstrates a convex reformulation of the RAPS optimization problems by converting transforming the mixed-binary variables into linear constraints, enabling efficient solution using modern optimization tools. We explore and evaluate two approaches to specifying performance: one involving the full information matrix and another focusing solely on its diagonal elements. Additionally, this paper includes a simulation comparison of the information and risk outcomes from KF using all measurement, TD for outlier removal, and RAPS for measurement selection.

II. PROBLEM STATEMENT

The discrete-time evolution of the linear system is

$$\mathbf{x}_k = \mathbf{F} \mathbf{x}_{k-1} + \mathbf{w}_k \quad (1)$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector at discrete-time $t_k = kT$ where T is the sampling interval, \mathbf{F} is the known transition matrix and $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ is white Gaussian process noise with covariance matrix \mathbf{Q} . The time propagation of the state vector \mathbf{x}_k and its covariance \mathbf{P}_k is assumed to be unaffected by outliers. Given the Gaussian probability density function of the posterior $\mathcal{N}(\hat{\mathbf{x}}_{k-1}^+, \mathbf{P}_{k-1}^+)$, the state estimate and its error covariance are propagated through time as

$$\hat{\mathbf{x}}_k^- \doteq \mathbf{F} \hat{\mathbf{x}}_{k-1}^+ \quad (2)$$

$$\mathbf{P}_k^- \doteq \mathbf{F} \mathbf{P}_{k-1}^+ \mathbf{F}^\top + \mathbf{Q} \quad (3)$$

The measurement vector at time t_k is modeled as:

$$\mathbf{y}_k = \mathbf{H} \mathbf{x}_k + \boldsymbol{\eta}_k \quad (4)$$

where $\mathbf{y}_k \in \mathbb{R}^m$, $\mathbf{H} \in \mathbb{R}^{m \times n}$ is the known measurement matrix, and $\boldsymbol{\eta}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$ represents the white Gaussian measurement noise. The measurement noise matrix $\mathbf{R}_k \in \mathbb{R}^{m \times m}$ is assumed to be diagonal and invertible. Therefore, it can be written as $\mathbf{R} = \sum_{i=1}^m \sigma_i^2 \mathbf{e}_i \mathbf{e}_i^\top$, where \mathbf{e}_i is the i^{th} standard basis vector.

Some measurements might be affected by outliers. Outlier measurements are those that are improbable relative to the model stated in eqn. (4). They may be described as

$$y_k(i) = \mathbf{h}_i \mathbf{x}_k + \eta_k(i) + s_i, \quad (5)$$

where $y_k(i)$ represents the i -th element of \mathbf{y}_k , that is affected by the outlier s_i , $\eta_k(i)$ is the i -th element of $\boldsymbol{\eta}_k$, and \mathbf{h}_i is the i -th row of \mathbf{H} . The outlier s_i is unknown and has no established model. The challenge lies in accommodating these outlier measurements during the measurement update process.

III. STATE ESTIMATION USING SELECTED MEASUREMENTS

This section reviews different strategies to accommodate outliers in the state estimation measurement update. Each strategy either selects or discards certain measurements. The discussion is facilitated by introducing a binary vector $\mathbf{b} = [b_1, \dots, b_m]^\top$. The binary element b_i indicates whether to include ($b_i = 1$) or exclude ($b_i = 0$) measurement $y_k(i)$ [22].

For a given value of \mathbf{b} , the maximum-a-posteriori (MAP) approach estimates the state by minimizing the cost function

$$\hat{\mathbf{x}}_k^+ = \underset{\mathbf{x}_k}{\operatorname{argmin}} C(\mathbf{x}_k, \mathbf{b}), \quad (6)$$

where the cost function (see [18]) is

$$\begin{aligned} C(\mathbf{x}_k, \mathbf{b}) &= \|\mathbf{x}_k - \hat{\mathbf{x}}_k^-\|_{\mathbf{P}_k^-}^2 + \|\boldsymbol{\Sigma}_\mathbf{R} \boldsymbol{\Phi}(\mathbf{b}) (\mathbf{y}_k - \mathbf{H} \mathbf{x}_k)\|^2 \\ &= \|\mathbf{x}_k - \hat{\mathbf{x}}_k^-\|_{\mathbf{P}_k^-}^2 + \sum_{i=1}^m \frac{b_i^2}{\sigma_i^2} (y_i - \mathbf{h}_i \mathbf{x}_k)^2 \end{aligned} \quad (7)$$

$$= \|\mathbf{x}_k - \hat{\mathbf{x}}_k^-\|_{\mathbf{P}_k^-}^2 + \sum_{i=1}^m \frac{b_i}{\sigma_i^2} (y_i - \mathbf{h}_i \mathbf{x}_k)^2 \quad (8)$$

where $\forall \mathbf{f} \in \mathbb{R}^\ell$, $\boldsymbol{\Phi}(\mathbf{f}) = \operatorname{diag}(\mathbf{f}) : \mathbb{R}^\ell \mapsto \mathbb{R}^{\ell \times \ell}$ and $b_i^2 = b_i$ due to the binary nature of b_i . The covariance matrices $\mathbf{R}^{-1} = \boldsymbol{\Sigma}_\mathbf{R}^\top \boldsymbol{\Sigma}_\mathbf{R}$ and $(\mathbf{P}_k^-)^{-1} = \boldsymbol{\Sigma}_\mathbf{P}^\top \boldsymbol{\Sigma}_\mathbf{P}$ are represented using the squared Mahalanobis distance (e.g., $\|\mathbf{x}_k - \hat{\mathbf{x}}_k^-\|_{\mathbf{P}_k^-}^2 = \|\boldsymbol{\Sigma}_\mathbf{P} (\mathbf{x}_k - \hat{\mathbf{x}}_k^-)\|^2$ see [23]).

The posterior information matrix is computed as

$$\begin{aligned} \mathbf{J}_k^+ &= \mathbf{H}^\top \boldsymbol{\Phi}(\mathbf{b})^\top \mathbf{R}_k^{-1} \boldsymbol{\Phi}(\mathbf{b}) \mathbf{H} + \mathbf{J}_k^- \\ &= \sum_{i=1}^m \frac{b_i}{\sigma_i^2} \mathbf{h}_i^\top \mathbf{h}_i + \mathbf{J}_k^- \end{aligned} \quad (9)$$

where $\mathbf{J}_k^- = (\mathbf{P}_k^-)^{-1}$ denotes the prior information matrix. The corresponding posterior covariance matrix is computed as $\mathbf{P}_k^+ = (\mathbf{J}_k^+)^{-1}$. Two approaches to calculate the measurement selection vector \mathbf{b} are considered below.

A. Threshold Decisions (TD)

The traditional method utilizes a threshold test to decide whether a measurement is compromised by an outlier:

$$b_i \doteq \begin{cases} 0, & \text{when } |r_i| \geq \lambda \sigma_{r_i} \\ 1, & \text{when } |r_i| < \lambda \sigma_{r_i} \end{cases} \quad (10)$$

where $\lambda > 0$ is the decision threshold, $\sigma_{r_i}^2 = \mathbf{h}_i \mathbf{P}^- \mathbf{h}_i^\top + \sigma_i^2$ is the covariance of residual component r_i (see e.g.: [12], [13]). The residuals are computed as

$$r_i \doteq y_i - \mathbf{h}_i \hat{\mathbf{x}}_k^-. \quad (11)$$

After \mathbf{b}_{TD} is determined by eqn. (10), the optimal estimate and information are the solutions of eqn. (6) and eqn. (9).

Many studies have critiqued the TD method for its failure to reliably detect outliers, pointing out its susceptibility to missed detections and false alarms [18], [24]. Missed detections can result in corrupted state estimates and overconfident covariance matrices, creating inconsistencies between actual and estimated uncertainties. This discrepancy can significantly impair the reliability of future subsequent outlier decisions. In addition, fixed decision threshold means that outlier decisions do not take into account any performance specifications.

B. Risk-Averse Optimization

The RAPS approach selects an optimal subset of measurements that to minimize the Bayesian risk as quantified by eqn. (8), subject to a specified performance criteria [18], [20]. Two constraint implementations are considered herein.

To facilitate the discussion, we define two symbols. Dropping the subscript k , starting from eqn. (9) the posterior information matrix computed for a specific choice of \mathbf{b} is

$$\mathbf{J}^+(b) = \sum_{i=1}^m \frac{b_i}{\sigma_i^2} \mathbf{h}_i^\top \mathbf{h}_i + \mathbf{J}^-. \quad (12)$$

1) *Full-RAPS*: This approach constrains the posterior information matrix within a semi-definite framework: $\mathbf{J}^+(b) \succcurlyeq \Phi(\mathbf{J}_d)$ where the symbol \mathbf{J}_d represents a user-defined non-negative performance lower bound vector. Defining $\mathbf{F}_i = \frac{1}{\sigma_i^2} \mathbf{h}_i^\top \mathbf{h}_i$, eqn. (12) becomes

$$\sum_{i=1}^m \mathbf{F}_i b_i \succcurlyeq \Phi(\mathbf{J}_d) - \mathbf{J}^-, \quad (13)$$

With this Linear Matrix Inequality (LMI), the Full-RAPS optimization problem is

$$\left. \begin{aligned} \mathbf{x}_k^+, \mathbf{b}^* &= \underset{\mathbf{x}_k, \mathbf{b}}{\operatorname{argmin}} C(\mathbf{x}_k, \mathbf{b}) \\ \text{s.t.: } \sum_{i=1}^m \mathbf{F}_i b_i &\succcurlyeq \Phi(\mathbf{J}_d) - \mathbf{J}^- \\ b_i &\in \{0, 1\} \text{ for } i = 1, \dots, m \end{aligned} \right\} \quad (14)$$

where the cost $C(\mathbf{x}_k, \mathbf{b})$ is defined in eqn. (8).

2) *Diag-RAPS*: This approach chooses the constraint $\operatorname{diag}(\mathbf{J}^+(b)) \geq \mathbf{J}_d$. By constraining only the diagonal elements of the posterior information matrix computational complexity is reduced relative to the semi-definiteness check in eqn. (13). Appendix A of [25], shows this constraint is equivalent to

$$\mathbf{G}\mathbf{b} \geq \mathbf{d} \text{ where } \mathbf{d} = \mathbf{J}_d - \operatorname{diag}(\mathbf{J}_k^-) \quad (15)$$

$$\text{and } \mathbf{G} = \begin{bmatrix} \frac{h_{11}^2}{\sigma_1^2} & \dots & \frac{h_{m1}^2}{\sigma_m^2} \\ \vdots & \ddots & \vdots \\ \frac{h_{1m}^2}{\sigma_1^2} & \dots & \frac{h_{mm}^2}{\sigma_m^2} \end{bmatrix}. \text{ With this linear constraint on } \mathbf{b},$$

the Diag-RAPS optimization problem is

$$\left. \begin{aligned} \mathbf{x}_k^+, \mathbf{b}^* &= \underset{\mathbf{x}_k, \mathbf{b}}{\operatorname{argmin}} C(\mathbf{x}_k, \mathbf{b}) \\ \text{s.t.: } \mathbf{G}\mathbf{b} &\geq \mathbf{d} \\ b_i &\in \{0, 1\} \text{ for } i = 1, \dots, m \end{aligned} \right\} \quad (16)$$

3) *Comparison of Full and Diagonal RAPS*: The cost function of eqn. (8) is quadratic in \mathbf{x}_k and linear in \mathbf{b} when these variables are treated separately; however, it is jointly non-convex. Therefore, Problems (14) and (16) are each mixed-integer non-convex problems. Given the binary nature of \mathbf{b} , [20] proposed exhaustive search methods with worst-case time complexity $O(m!)$. For each choice of \mathbf{b} , the optimization which is quadratic in \mathbf{x}_k is straightforward.

The search over \mathbf{b} could be implemented using an m -layer binary tree search, leveraging the standard branch-and-bound (BNB) with worst-case time complexity to $O(2^m)$.

IV. CONVEX REFORMULATION

There now exist software packages to solve convex mixed-integer programming problems; however, Problems (14) and (16) are characterized by non-convex objective functions coupled with convex constraints. This section transforms the non-convex optimization cost function into a convex format, thereby transforming Problems (14) and (16) into convex problems.

The cost function in eqn. (7) is equivalent to

$$C(\mathbf{x}_k, \mathbf{b}) = \|\mathbf{x}_k - \hat{\mathbf{x}}_k^-\|_{\mathbf{P}_k}^2 + \sum_{i=1}^m \frac{1}{\sigma_i^2} (y_i b_i - \mathbf{h}_i \mathbf{b}_i \mathbf{x})^2. \quad (17)$$

Let $\mathbf{q} = [\mathbf{q}_1; \dots; \mathbf{q}_i; \dots; \mathbf{q}_m] \in \mathbb{R}^{n \times m}$, where $\mathbf{q}_i = b_i \mathbf{x} = [q_{i1}, \dots, q_{ij}, \dots, q_{in}]^\top \in \mathbb{R}^n$ with $q_{ij} = b_i x_j$. Although the product $b_i x_j$ is inherently non-convex and mixed-binary, it can be encapsulated within a linear constraint set:

$$\mathcal{A}(\mathbf{x}_k, \mathbf{b}, \mathbf{q}) = \begin{cases} \mathcal{A}_1 : \mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max}, \\ \mathcal{A}_2 : \mathbf{x} - \mathbf{x}_{\max}(1 - b_i) \leq \mathbf{q}_i \leq b_i \mathbf{x}_{\max}, \\ \mathcal{A}_3 : b_i \mathbf{x}_{\min} \leq \mathbf{q}_i \leq \mathbf{x} - \mathbf{x}_{\min}(1 - b_i), \end{cases} \quad (18)$$

where \mathbf{x}_{\min} and \mathbf{x}_{\max} denote lower and upper bounds on \mathbf{x} , respectively. For state estimation problems, the posterior state estimate is typically within a limited distance to the prior state. Therefore, these bounds can be configured as $\hat{\mathbf{x}}^- \pm \Delta$, with Δ being a suitably large positive value.

The linear constraints in eqn. \mathcal{A}_2 and eqn. \mathcal{A}_3 establish important properties for \mathbf{q} :

- 1) For $b_i = 1$, \mathcal{A}_2 simplifies to $\mathbf{x} \leq \mathbf{q}_i \leq \mathbf{x}_{\max}$, and \mathcal{A}_3 simplifies to $\mathbf{x}_{\min} \leq \mathbf{q}_i \leq \mathbf{x}$, ensuring $\mathbf{q}_i = \mathbf{x}$.
- 2) For $b_i = 0$, \mathcal{A}_2 becomes $\mathbf{x} - \mathbf{x}_{\max} \leq \mathbf{q}_i \leq 0$, and \mathcal{A}_3 becomes $0 \leq \mathbf{q}_i \leq \mathbf{x} - \mathbf{x}_{\min}$, ensuring $\mathbf{q}_i = \mathbf{0}$.

These linear (convex) constraints therefore ensure $q_{ij} = b_i x_j$. Note that the binary vector \mathbf{b} cannot be relaxed to a non-binary variable, as that would compromise the require property that $\mathbf{q}_i = b_i \mathbf{x}$.

For binary \mathbf{b} , with the constraints in eqns. (18), the cost function in eqn. (7) has the convex form:

$$C_c(\mathbf{x}_k, \mathbf{b}, \mathbf{q}) = \|\mathbf{x}_k - \hat{\mathbf{x}}_k^-\|_{\mathbf{P}_k}^2 + \sum_{i=1}^m \frac{(y_i b_i - \mathbf{h}_i \mathbf{q}_i)^2}{\sigma_i^2}.$$

Both Problems (14) and (16)) can now be written in convex form. RAPS Optimization using Full Information matrix (Full-RAPS) is equivalent to

$$\left. \begin{aligned} \mathbf{x}_k^+, \mathbf{b}^*, \mathbf{q} &= \underset{\mathbf{x}_k, \mathbf{b}, \mathbf{q}}{\operatorname{argmin}} C_c(\mathbf{x}_k, \mathbf{b}, \mathbf{q}) \\ \text{s.t.: } \sum_{i=1}^m \mathbf{F}_i b_i &\succcurlyeq (\mathbf{J}_d) - \mathbf{J}_k^- \\ b_i &\in \{0, 1\} \text{ for } i = 1, \dots, m \\ \mathcal{A}(\mathbf{x}_k, \mathbf{b}, \mathbf{q}), \end{aligned} \right\} \quad (19)$$

RAPS Optimization using Diagonal Information matrix (Diag-RAPS) is equivalent to

$$\left. \begin{aligned} \mathbf{x}_k^+, \mathbf{b}^*, \mathbf{q} &= \underset{\mathbf{x}_k, \mathbf{b}, \mathbf{q}}{\operatorname{argmin}} C_c(\mathbf{x}_k, \mathbf{b}, \mathbf{q}) \\ \text{s.t.: } \mathbf{G}\mathbf{b} &\geq \mathbf{d} \\ b_i &\in \{0, 1\} \text{ for } i = 1, \dots, m \\ \mathcal{A}(\mathbf{x}_k, \mathbf{b}, \mathbf{q}). \end{aligned} \right\} \quad (20)$$

Both Problem (19) and Problem (20) are convex mixed-integer optimization problems featuring nonlinear objective functions. Problem (20) incorporates linear constraints. In contrast, Problem (19), not only has linear constraints, but also involves a LMI constraint, adding a layer of complexity to its formulation. Both involve binary variables.

Both problems are solvable using the BNB algorithm in YALMIP [26]. While in the worst case the BNB algorithm exhibits exponential time complexity, on average it is more efficient than the previously proposed DFS strategy. In principle, both Problem (19) and Problem (20) could be solved more efficiently by combining the BNB algorithm with convex optimization. However, contemporary solvers like MOSEK and Gurobi [27], [28], that are commonly used in the CVX [29] and YALMIP optimization environments, do not currently support integer variables with semi-definite matrix inequalities. Therefore, MOSEK and Gurobi are suitable for Problem (20), but not Problem (19). A comparative analysis of computation times is presented in Sec. V-D, using YALMIP BNB for Full-RAPS and CVX MOSEK for Diag-RAPS.

V. EVALUATION IN SIMULATION

This section evaluates and compares Full-RAPS and Diag-RAPS on computation time. Then, state estimation performance is evaluated for: the KF without outlier removal, the KF equipped with TD outlier removal as discussed in Section III-A with $\lambda = 2$, and Diag-RAPS solved using MOSEK. The performance evaluation was conducted in MATLAB on an AMD 5800X CPU.

A. System Model

The systems that is simulated has the state vector

$$\mathbf{x} = [\mathbf{p} \quad \mathbf{v} \quad \mathbf{a}]^T \in \mathbb{R}^9$$

where $\mathbf{p} \in \mathbb{R}^3$ represents position (components denoted as N , E , and D), $\mathbf{v} \in \mathbb{R}^3$ represents velocity, and $\mathbf{a} \in \mathbb{R}^3$ represents acceleration. Therefore, $n = 9$. During the simulation, the acceleration signal is selected to cause the vehicle to follow a trajectory similar to a vehicle driving around a city block that is square with each edge being 200 meters long. The origin of the coordinate system is at the center of this square. The D component undulates in a slow sinusoidal fashion as the vehicle traces its trajectory.

The navigation system does not know the acceleration of the vehicle. For the purpose of state estimation, the differential equations for this system are

$$\dot{\mathbf{p}}(t) = \mathbf{v}(t), \quad \dot{\mathbf{v}}(t) = \mathbf{a}(t), \quad \text{and} \quad \dot{\mathbf{a}}(t) = \boldsymbol{\omega}(t) \quad (21)$$

where $\boldsymbol{\omega}(t)$ is continuous-time white Gaussian process noise with power spectral density \mathbf{S} . This is referred to a position-velocity-acceleration model and is frequently used in GNSS applications [25], [30].

To estimate the state vector, each state estimator uses a vector of measurements that occurs every $T = 1$ second. Each estimator is designed based on the model in eqn. (4). Each row of the measurement matrix has the form

$$\mathbf{h}_i = [h_{iN} \quad h_{iE} \quad h_{iD} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0], \quad (22)$$

where $\|\mathbf{h}_i\|_2 = 1$ and $h_{iz} > 0$. These first three components, related to the position vector, represent the line-of-sight unit vector pointing from the receiver antenna to the satellite (see Section 8.2 in [31]). The simulation considers m satellites, with randomly distributed azimuth and elevation, that are fixed during the duration of the simulation.

B. Measurements and Outliers

The simulated measurements are produced using the model in eqn. (5), where $\operatorname{var}(\eta_i) = \sigma_i^2$ with $\sigma_i = 1.5$ m. The ground truth trajectory and outlier corrupted measurements were created once and then used by all three estimators.

Each measurement outlier signal $s_i(kT)$ was defined to have properties similar to multipath in GNSS. Dropping the i , because the model is the same for all measurements, the variance of each outlier is

$$\sigma_s^2 = \sigma_{el}^2 + \sigma_{az}^2,$$

where $\sigma_{el} = \frac{0.6}{|\psi| + \epsilon}$ and $\sigma_{az} = \frac{0.3}{\theta + \epsilon}$. In this expression, $\theta > 0$ and ψ represent the radian elevation and azimuth angles of the satellite relative to the receiver and ϵ is a small positive number to prevent division by zero. For the i -th measurement, at each time $t_k = kT$, an outlier $s_i(kT)$ is computed as a zero mean Gaussian random variable with standard deviation $\sigma_{s_i}^2$ based on θ_i and ψ_i .

Because the number of reflecting surfaces along the signal path that can cause multipath increases as elevation decreases, the outlier variance is designed to increase rapidly as the satellite elevation decreases. The variance of the outlier with respect to azimuth is designed to correspond to a building or other reflecting surface near the origin of the simulated system that causes high multipath. Because the vehicle trajectory goes around the building the specific direction corresponding to the high variance changes with the vehicle location.

C. Performance Specification

The performance specification \mathbf{J}_d , defines a the lower bound in the constraint in Problems (19) and (20). It was selected to only constrain the information required for the three position variables. The Society of Automotive Engineers (SAE) specification for positioning accuracy in highway vehicle applications requires the horizontal accuracy better than 1.5 meters and vertical accuracy 3 meters, both at a 68% probability level. The determination of \mathbf{J}_d based on the SAE criteria is discussed in Section 7.2 of [25], yielding

$$\mathbf{J}_d = [1.389, 1.389, 0.347].$$

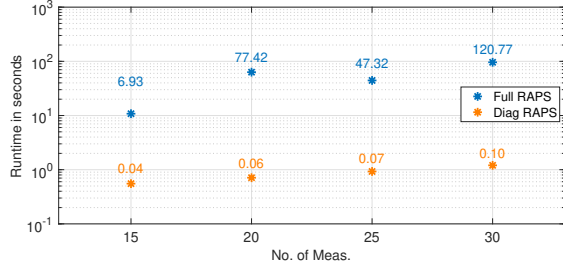


Fig. 1: Runtime comparison between Full-RAPS (blue) vs Diag-RAPS (orange). Each point represents the mean of the per epoch runtime for an experiment with the number of measurements specified along the horizontal axis. The values above the points represent the STD.

Additional states could be constrained, but this constraint is appropriate for this application. If constraints were desired for velocity (or acceleration) which are not observable until position measurements are available for at least two (three) distinct times, then the lower bound would need to change with time.

D. Runtime comparison between Full-RAPS and Diag-RAPS

This section evaluates the real-time feasibility of the Full-RAPS and Diag-RAPS approaches. Four experiments were conducted, each utilizing a fixed number of 15, 20, 25, and 30 measurements per epoch. The Full-RAPS measurement update corresponds to the solutions of the optimization problems defined in Problem (19), solved using the YALMIP BNB toolbox. The Diag-RAPS update pertains to the solutions of the optimization problems defined in Problem (20), solved via CVX MOSEK toolbox. The exhaustive search strategy in [20] was not included in this study due to the prohibitive computational cost (e.g., more than 10 minutes for a single epoch with $m = 15$), especially as m increases.

As m increases to 20, the Full-RAPS runtime extends beyond one minute per measurement epoch, rendering the evaluation of the complete dataset impractical. Consequently, the comparison focuses on a subset of 20 measurement epochs. Fig. 1 compares the computational times for each method as a function of the number of measurements. The average computation time trends upward with the number of measurements. Diag-RAPS achieves a significantly lower computational cost, approximately 100 times faster than Full-RAPS, and exhibits a reduced Standard Deviation (STD). Based on these findings, only Diag-RAPS was investigated for state estimation performance comparisons.

E. Estimator Implementations

The time propagation for all three estimators is identical, using the state and covariance propagation stated in eqns. (2) and (3). The computation of the parameters \mathbf{F} and \mathbf{Q} for the discrete-time model discussed in Section 4.7 of [31] and derived in Appendix B of [25].

In this section, the number of measurements is $m = 50$. Each estimator implements the measurement update differ-

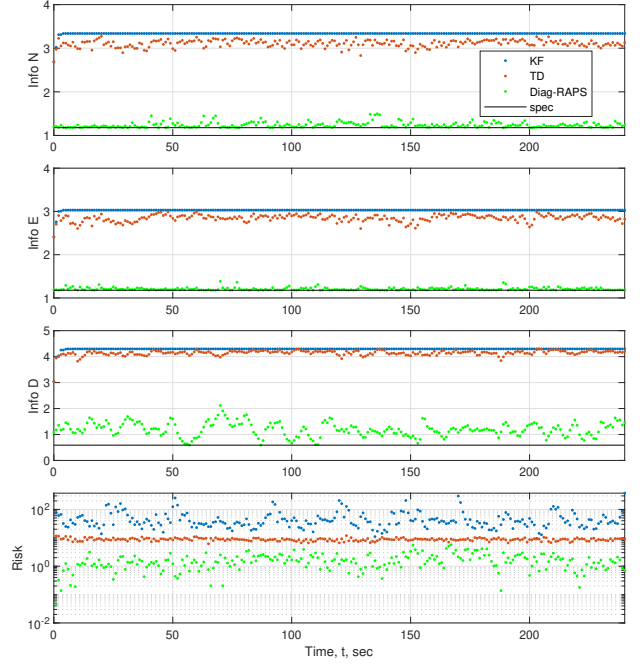


Fig. 2: Simulation results comparing the square root of the posterior information and of the risk for Kalman filter (KF) in blue, Kalman filter with threshold detection (TD) in red, and Diag-RAPS in green. The solid black line shows the corresponding square root of the diagonal element of J_d .

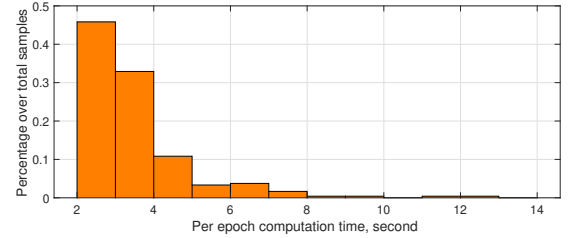


Fig. 3: Diag-RAPS computation time probability histogram.

ently resulting in a different risk as quantified by a cost equivalent to eqn. (17).

- KF: The standard KF uses the \mathbf{x} that minimizes the MAP cost in eqn. (6) for $\mathbf{b} = \mathbf{1}$.
- TD: The KF with TD uses the \mathbf{x} that minimizes the MAP cost in eqn. (6) with $\mathbf{b} = \mathbf{b}_{TD}$ (see eqn. (10)).
- RAPS: Diag-RAPS uses the \mathbf{x} that solves Problem 20.

F. Results

The estimation accuracies for the three estimators are shown in Fig. 2. The top three subplots show the square root of diagonal elements of the posterior information matrix, which correspond to the north, east, and down position states. The bottom subplot shows the square root of the risk incurred by each estimator.

Each plot of the square root of the posterior information shows: the information achieved by the KF as blue dots; the information achieved by the TD as red dots; the information

achieved by Diag-RAPS as green dots; the specified level of performance (i.e., square root of the corresponding component of \mathbf{J}_d) as a solid black line. The same color scheme is used for the plot of the square root of the risk.

The KF graphs always achieves the highest (predicted) information for all components, but also the highest risk, because it uses all measurements. The TD graphs show that, by ignoring those measurements whose normalized residual fails the threshold test, it is able to reduce the risk. The trade-off is that less (predicted) information is extracted from the measurements. The Diag-RAPS graphs show that it consistently maintains the lowest risk while always achieving the specified performance levels. It is critical to understand that a high-level of predicted information is not likely to be achieved when the outlier risk is also high.

In terms of computational efficiency, both KF and TD methods require only a few milliseconds per measurement epoch, as they do not involve real-time optimization problems. Fig. 3 presents a histogram showing the percentage of computation times (per measurement epoch) within each bin for Diag-RAPS over 240 epochs, showing that it typically resolves within a few seconds.

VI. CONCLUSION AND FUTURE RESEARCH

The main contributions of this article are the presentation and demonstration of a convex reformulation of the RAPS state estimation problem. The article applied this novel approach to two variations of RAPS to attain their globally optimal solution. The Full-RAPS approach imposes constraints on the full information matrix, whereas the Diag-RAPS approach constrains the diagonal elements of the information matrix. Both of them yield a mixed-integer convex programming problem that can be resolved using existing software optimization tools. Our simulation evaluations demonstrate that Diag-RAPS is approximately 100 times faster than Full-RAPS, making it a more practical choice for real-world applications. The state estimation performance of Diag-RAPS is benchmarked against a standard KF using all measurements and the TD approach based on checking normalized residual magnitudes against a threshold. The convex reformulation has proven effective, with simulation outcomes indicating that the RAPS solution meets the defined performance specifications while minimizing risk.

Regarding future research, this paper has concentrated on the solvability aspects of the RAPS problem. A very interesting next step is to study the actual positioning performance of Diag-RAPS using real-world data. This necessitates reducing computation costs to facilitate real-time application feasibility [32]. On alternative approach to drastically improving computation time is based on the idea that the globally optimal measurement selection is not required; instead, what is required is a feasible solution that is not high risk [18].

VII. ACKNOWLEDGMENT

The authors gratefully acknowledge the UCR KA Endowment and the US DOT CARNATIONS Center, each of which has partially funded this research. The ideas reported herein,

and any errors or omissions, are the responsibility of the authors and do not reflect the opinions of the sponsors.

REFERENCES

- [1] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements and Performance Second Edition*. Ganga-Jamuna Press, 2006.
- [2] S. Camacho-Lara, "Current and future GNSS and their augmentation systems," *Handbook of Satellite Applications*, p. 781, 2017.
- [3] R. Odolinski, H. Yang, L.-T. Hsu, M. Khider, G. M. Fu, and D. Dusha, "Evaluation of the Multi-GNSS, Dual-Frequency RTK Positioning Performance for Recent Android Smartphone Models in a Phone-to-Phone Setup," in *ITM 2024*, 2024, pp. 42–53.
- [4] A. Ramanandan, A. Chen, and J. A. Farrell, "Observability Analysis of INS and Lever-Arm Error States with CDGPS-Camera Aiding," in *IEEE/ION PLANS*, 2010, pp. 1197–1203.
- [5] S. Su, P. Zhu, and W. Ren, "An Optimization Approach to Fully Distributed Active Joint Localization and Target Tracking in Multi-Robot Systems," in *2022 American Control Conference (ACC)*, 2022, pp. 5081–5086.
- [6] Y. Song, "Deep Learning Applications in the Medical Image Recognition," *American Journal of Computer Science and Technology*, vol. 2, no. 2, pp. 22–26, 2019.
- [7] Y. Kang, Y. Xu, C. P. Chen, G. Li, and Z. Cheng, "6: Simultaneous tracking, tagging and mapping for augmented reality," in *SID Symposium Digest of Technical Papers*, vol. 52. Wiley Online Library, 2021, pp. 31–33.
- [8] P. Li, Y. Lin, and E. Schultz-Fellenz, "Contextual hourglass network for semantic segmentation of high resolution aerial imagery," *arXiv preprint arXiv:1810.12813*, 2019.
- [9] T. Deng, N. Wang, C. Wang, S. Yuan *et al.*, "Incremental joint learning of depth, pose and implicit scene representation on monocular camera in large-scale scenes," *arXiv preprint arXiv:2404.06050*, 2024.
- [10] T. Deng, H. Xie, J. Wang, and W. Chen, "Long-term visual simultaneous localization and mapping: Using a bayesian persistence filter-based global map prediction," *IEEE Robotics & Automation Magazine*, vol. 30, no. 1, pp. 36–49, 2023.
- [11] V. Barnett and T. Lewis, *Outliers in Statistical Data*. Wiley, 1994.
- [12] P. M. Frank and X. Ding, "Survey of Robust Residual Generation and Evaluation Methods in Observer-Based Fault Detection Systems," *J. of Process Control*, vol. 7, pp. 403–424, 1997.
- [13] R. J. Patton and J. Chen, "A review of parity space approaches to fault diagnosis," *IFAC Proceedings Volumes*, vol. 24, no. 6, pp. 65–81, 1991.
- [14] P. J. Rousseeuw and K. Van Driessen, "Computing LTS regression for large data sets," *Data mining and knowledge discovery*, vol. 12, no. 1, pp. 29–45, 2006.
- [15] R. A. Maronna, R. D. Martin, V. J. Yohai, and M. Salibián-Barrera, *Robust Statistics: Theory and Methods*. John Wiley & Sons, 2019.
- [16] M. Zhu, Y. Zhang, Y. Gong, K. Xing, X. Yan, and J. Song, "Ensemble methodology: Innovations in credit default prediction using lightgbm, xgboost, and localensemble," *arXiv preprint arXiv:2402.17979*, 2024.
- [17] P. Li, M. Abouelenien, R. Mihalcea, Z. Ding, Q. Yang, and Y. Zhou, "Deception detection from linguistic and physiological data streams using bimodal convolutional neural networks," *arXiv preprint arXiv:2311.10944*, 2023.
- [18] W. Hu, J.-B. Uwineza, and J. A. Farrell, "Outlier Accommodation for GNSS Precise Point Positioning using Risk-Averse State Estimation," *arXiv preprint arXiv:2402.01860*, 2024.
- [19] W. Hu, Y. Hu, M. Stas, and J. A. Farrell, "Optimization-Based Outlier Accommodation for Tightly Coupled RTK-Aided Inertial Navigation Systems in Urban Environments," *arXiv preprint arXiv:2407.13912*, 2024.
- [20] E. Aghapour, F. Rahman, and J. A. Farrell, "Outlier accommodation in nonlinear state estimation: A risk-averse performance-specified approach," *IEEE T-CST*, pp. 1–13, 2019.
- [21] Y. Zhang, M. Zhu, K. Gui, J. Yu, Y. Hao, and H. Sun, "Development and application of a monte carlo tree search algorithm for simulating da vinci code game strategies," *arXiv preprint arXiv:2403.10720*, 2024.
- [22] L. Carlone, A. Censi, and F. Dellaert, "Selecting Good Measurements via ℓ_1 Relaxation: A Convex Approach for Robust Estimation over Graphs," in *2014 IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems*, 2014, pp. 2667–2674.

- [23] F. Dellaert and M. Kaess, "Square Root SAM: Simultaneous Localization and Mapping via Square Root Information Smoothing," *The International Journal of Robotics Research*, vol. 25, no. 12, pp. 1181–1203, Dec. 2006. [Online]. Available: <http://journals.sagepub.com/doi/10.1177/0278364906072768>
- [24] E. Aghapour and J. A. Farrell, "Outlier Accommodation in Sensor Rich Environments: Moving Horizon Risk-Averse Performance-Specified State Estimation," in *58th CDC*. IEEE, 2019, pp. 7917–7922.
- [25] W. Hu, "Optimization-Based Risk-Averse Outlier Accommodation With Linear Performance Constraints: Real-Time Computation and Constraint Feasibility in CAV State Estimation," *UC Riverside*, 2024. [Online]. Available: <https://escholarship.org/uc/item/5mf54016>
- [26] J. Löfberg, "YALMIP : A Toolbox for Modeling and Optimization in MATLAB," in *In Proc. of the CACSD Conf.*, Taipei, Taiwan, 2004.
- [27] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 10.1.28*, 2019. [Online]. Available: <http://docs.mosek.com/9.0/toolbox/index.html>
- [28] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," 2023. [Online]. Available: <https://www.gurobi.com>
- [29] M. Grant and S. Boyd, "CVX Users' Guide, Release 2.2," *online: https://cvxr.com/cvx/doc/CVX.pdf*, 2020.
- [30] F. Rahman, E. Aghapour, and J. A. Farrell, "Vehicle ECEF Position Accuracy and Reliability in the Presence of DGNSS Communication Latency," *IEEE Intell. Transp. Syst. Mag.*, pp. 0–0, 2020.
- [31] J. A. Farrell, *Aided Navigation: GPS with High Rate Sensors*. McGraw-Hill, Inc., 2008.
- [32] Z. Ding, Z. Lai, S. Li, P. Li, Q. Yang, and E. Wong, "Confidence trigger detection: Accelerating real-time tracking-by-detection systems," *arXiv preprint arXiv:1902.00615*, 2024.