Kalman Filtering for Indoor Pose Estimation on Riemannian Manifolds Using IMU and Acoustics

Mohammed H. AlSharif*, Mohanad Ahmed*, Mohamed Siala§, Tareq Y. Al-Naffouri¶
*Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, USA
§Department of Applied Mathematics, Signals and Communication, SUP'COM, Ariana, Tunisia
¶Electrical and Computer Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
*{alsharmh, mohanad}@caltech.edu, §mohamed.siala@supcom.tn, ¶tareq.alnaffouri@kaust.edu.sa

***Abstract*—This paper presents a novel approach that combines inertial navigation systems (INS) with an acoustic Riemannian-based localization system to enhance indoor pose tracking. The proposed method employs an Extended Kalman Filter (EKF) to fuse data from the two systems. The Riemannian-based localization system delivers high-accuracy estimates of the target's pose, which is then used to correct the INS data. A new projection algorithm is introduced to map the EKF output onto the Riemannian manifold, further improving the accuracy of the estimation. Our results show that the proposed method significantly outperforms benchmark algorithms. The effectiveness of the proposed method was evaluated through our in-house experimental setup.**

***Index Terms*—Pose estimation, Inertial navigation, Acoustics.**

I. Introduction

Accurate pose estimation is crucial for various applications such as inventory management [1], smart homes, smart manufacturing, and health monitoring [2], [3]. High accuracy is fundamental to achieving the desired performance in these applications. Advances in sensor technology and computational techniques have led to the increased use of inertial sensors for navigation. Historically, inertial sensors were mechanically isolated from the target's rotational motion, ensuring high accuracy [4]. Modern inertial systems have simplified the mechanical complexity by rigidly attaching sensors to the target. These attached 'strapped down' sensors reduced cost and size, but increased computational complexity. In this paper, we focus on a strap-down inertial navigation system (INS) to estimate position, orientation, and velocity.

The INS operates based on kinematics and classical Newtonian mechanics, utilizing inertial measurement units (IMUs) comprised of accelerometers (motion sensors) and gyroscopes (rotation sensors) to measure acceleration and angular velocity. Although the INS can use these parameters to estimate position, orientation, and velocity at a high rate without external references, e.g. beacons, positioning, and orientation errors accumulate over time [2]. To mitigate these errors, navigation systems typically integrate additional positioning methods. In our previous work [5], we developed an acoustic-based pose estimation system using Riemannian optimization, which performs well under line-of-sight (LOS) conditions but suffers in non-line-of-sight (NLOS) scenarios due to the nature of the acoustic waves. Moreover, the system in [5] has a lower update rate compared to INS.

To address these limitations, we propose a Kalman filtering approach to integrate estimates from Riemannian localization methods [5] with those from the INS. This integration aims to enhance accuracy of the estimated pose in NLOS conditions and over extended periods.

The main contributions of this paper, detailed in Section IV, are as follows:

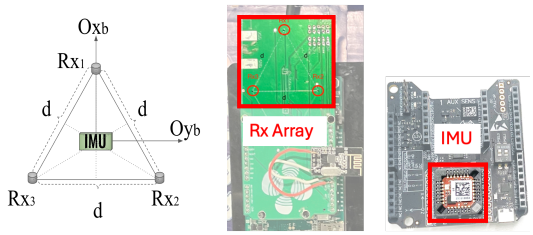
- A novel orientation correction method that models the orientation matrix as a state vector alongside position and velocity, resulting in new EKF formulations.
- A new projection algorithm to improve the accuracy of pose estimates obtained using EKF.
- Development and implementation of a new experimental setup to validate the proposed algorithms.
- An experimental evaluation of the proposed algorithms.

II. Problem Formulation

Accurate pose estimation of a moving target is pivotal in INS. While inertial navigation based on IMUs has high accuracy over short periods, its accuracy severely degrades over time due to error accumulation from the integration operations used in INS [4]. Therefore, it is common practice to fuse measurements from different sensors to improve the tracking accuracy of INS. We propose using three acoustic receivers arranged at the vertices of an isosceles triangle to improve INS-based pose estimation accuracy.

The proposed system consists of three acoustic receivers with unknown positions placed at the vertices of an isosceles triangle, four acoustic transmitters (beacons) with known positions, and an IMU placed at the centroid of the isosceles triangle. The three receivers, together with the IMU, form the mobile device (MD). While the proposed system is evaluated using an equilateral triangle, which is a special case of an isosceles triangle, the proposed algorithm can be applied to any isosceles triangle. The MD is illustrated in Figure 1.

Unlike INS, the three acoustic receivers provide highly accurate, non-degrading pose estimation for the MD but require LOS and operate at a slower rate. The proposed algorithm fuses both systems to achieve high-accuracy, high-rate pose estimation, even in NLOS conditions and over extended durations. The centroid position, velocity, and orientation of the triangle in Figure 1 are estimated using IMU measurements



(a) Illustration of MD (b) Acoustic Rx array with IMU

Fig. 1: Mobile device

and INS equations [4], and then corrected via our Riemannian localization [4].

III. Pose Estimation

This section provides a brief summary of the inertial navigation [4] and Riemannian localization algorithms [5] used in this work. For more details on INS and Riemannian localization, we refer the reader to [4] and [5].

A. Pose Estimation via Inertial Navigation System

Navigation systems must provide the pose of an MD relative to a fixed navigation frame, while the IMU measures acceleration and angular velocity in the MD's moving body frame. The body frame, centered at the triangle's centroid in Figure 1, aligns with the MD's roll, pitch, and yaw axes, whereas the navigation frame is fixed with axes north, east, and up. Both frames form right-hand systems.

Let's denote the navigation (reference) frame and body frame axes as $[O_{x_n}, O_{y_n}, O_{z_n}]$ and $[O_{x_b}, O_{y_b}, O_{z_b}]$, respectively. A vector $\mathbf{v}^b$ expressed in the body frame is transformed to the navigation frame via the orientation matrix $\mathbf{C}_b^n$ [4]:

$$\mathbf{v}^n = \mathbf{C}_b^n \mathbf{v}^b, \tag{1}$$

where $\mathbf{C}_b^n$ is a 3×3 matrix that represents the body frame's attitude relative to the navigation frame. This attitude can be expressed using direction cosines, Euler angles, or quaternions.

The IMU, equipped with a triaxial accelerometer and gyroscope, measures the specific force $\mathbf{a}^b = [a_x^b, a_y^b, a_z^b]^T$ (commonly referred to as 3D acceleration) and angular velocity $\omega_{nb}^b = [\omega_x, \omega_y, \omega_z]^T$ where the angular velocity is that of the body frame relative to the navigation frame, expressed in the body frame. To estimate the MD's position and velocity in the navigation frame under the assumption of known initial conditions, the orientation matrix $\mathbf{C}_b^n$ is computed, by integrating angular velocity over time [4]:

$$\mathbf{C}_b^n(k+1) = \mathbf{C}_b^n(k) + \Delta t \dot{\mathbf{C}}_b^n(k), \tag{2}$$

where Δt is the IMU sampling period, and $\dot{\mathbf{C}}_b^n(k) = \mathbf{C}_b^n(k)\mathbf{\Omega}_{nb}^b(k)$, and $\mathbf{\Omega}_{nb}^b$ is the skew-symmetric matrix of ω_{nb}^b. The orientation matrix is orthogonal, satisfying $\mathbf{C}_n^b(k) = (\mathbf{C}_b^n(k))^T$. The orthogonality of the orientation matrix is approximately achieved, assuming that the first-order approximation given in (2) is accurate under a high sampling rate.

The specific force $\mathbf{a}^b$ is transformed to the navigation frame as $\mathbf{a}^{n'}(k) = \mathbf{C}_b^n(k)\mathbf{a}^b(k)$, with gravitational acceleration g removed to obtain $\mathbf{a}^n(k)$. Assuming constant acceleration during the sampling period, the velocity $\mathbf{v}_c^n$ and position $\mathbf{p}_c^n$ at the centroid of the MD are updated as:

$$\mathbf{v}_c^n(k+1) = \mathbf{v}_c^n(k) + \Delta t\, \mathbf{a}^n(k),$$
$$\mathbf{p}_c^n(k+1) = \mathbf{p}_c^n(k) + \Delta t\, \mathbf{v}_c^n(k) + \frac{\Delta t^2}{2}\mathbf{a}^n(k).$$

Given initial conditions $\mathbf{p}_c^n(0)$, $\mathbf{v}_c^n(0)$, and $\mathbf{C}_b^n(0)$, the IMU measurements enable estimation of the MD's position, velocity, and orientation at each time step.

B. Pose Estimation via Acoustics on Riemannian Manifolds

The Riemannian optimization algorithm from [5] estimates the positions of three receivers arranged as vertices of an equilateral triangle with side length d, rigidly attached to the MD. The system includes four acoustic beacons with known positions. The MD's position is defined as the centroid of the triangle, and its orientation is derived from the vertices' positions.

Let $\mathbf{p}_i \in \mathbb{R}^3$ denote the i-th receiver's unknown position and $\mathbf{b}_j \in \mathbb{R}^3$ the j-th beacon's known position. The time of flight (TOF) τ_{ij} between the j-th beacon and i-th receiver, estimated using correlation-based methods [6] [7], is used to calculate distances $r_{ij} = c\tau_{ij}$, where c is the speed of sound. The receivers and beacons are assumed to be synchronized, as will be detailed in the experimental setup section. The position estimation problem is formulated as a constrained optimization [5]:

$$\min_{\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3 \in \mathbb{R}^3} \sum_{i=1}^{3} \left|\left| \mathbf{B}\mathbf{p}_i - \frac{1}{2}||\mathbf{p}_i||_2^2 \mathbf{1}_4 - \mathbf{y}_i \right|\right|_2^2 \tag{3a}$$

$$\text{s.t. } (\mathbf{p}_1 - \mathbf{p}_2)^T(\mathbf{p}_2 - \mathbf{p}_3) = -d^2\cos(\frac{\pi}{3}) \tag{3b}$$

$$(\mathbf{p}_1 - \mathbf{p}_3)^T(\mathbf{p}_2 - \mathbf{p}_3) = d^2\cos(\frac{\pi}{3}), \tag{3c}$$

where $\mathbf{B} \in \mathbb{R}^{4\times 3}$ is a matrix where the j-th row of the matrix corresponds to the location of the j-th beacon, i.e., $\mathbf{b}_j^T$, $\mathbf{1}_4$ is an all-ones vector, $\mathbf{y}_i = \frac{1}{2}(\mathbf{b}^2 - \mathbf{r}_i^2)$, $\mathbf{b}^2 \in \mathbb{R}^{4\times 1}$ is vector of squared norms of beacons positions, and $\mathbf{r}_i^2 \in \mathbb{R}^{4\times 1}$ is a vector of squared distances between the i-th receiver and each beacon. The constraints ensure the triangle's geometry, forming a Riemannian manifold $\mathcal{M}$.

Our algorithm in [5] estimates $\mathbf{p}_1, \mathbf{p}_2$, and $\mathbf{p}_3$ by solving the optimization problem in (3a) on Riemannian manifolds. The centroid position $\mathbf{p}_c^R$ and orientation matrix $\mathbf{C}_b^n$ are computed as:

$$\mathbf{p}_c^R = \sum_{i=1}^{3} \frac{\mathbf{p}_i}{3}, \quad \mathbf{C}_b^n = [O_{x_b}^n, O_{y_b}^n, O_{z_b}^n], \tag{4}$$

where the superscript R in $\mathbf{p}_c^R$ distinguishes the centroid position obtained using the Riemannian algorithm from the position obtained using an INS, and the body frame axes are:

$$O_{x_b}^n = \frac{\mathbf{p}_1 - \mathbf{p}_c^R}{||\mathbf{p}_1 - \mathbf{p}_c^R||} \tag{5}$$

$$O_{y_b}^{n'} = \mathbf{p}_2 - \mathbf{p}_c^R - < \mathbf{p}_2 - \mathbf{p}_c^R, O_{x_b}^n > O_{x_b}^n \tag{6}$$

$$O_{y_b}^n = \frac{O_{y_b}^{n'}}{||O_{y_b}^{n'}||}, \quad O_z^n = O_{x_b}^n \times O_{y_b}^n. \tag{7}$$

The vertices' positions are derived from the centroid and orientation matrix:

$$\mathbf{P} = [\mathbf{p}_c^R, \mathbf{p}_c^R, \mathbf{p}_c^R] + \mathbf{C}_b^n \begin{bmatrix} \frac{\sqrt{3}d}{3} & \frac{-\sqrt{3}d}{6} & \frac{-\sqrt{3}d}{6} \\ 0 & \frac{d}{2} & -\frac{d}{2} \\ 0 & 0 & 0 \end{bmatrix}. \quad (8)$$

IV. Fusion of INS and Acoustics Positioning

This section presents a new state equation formulation that incorporates the orientation matrix into the state vector, leading to revised EKF equations. A novel retraction operator is also introduced, projecting the EKF output onto the isosceles triangle manifold to improve tracking accuracy for the target's position and orientation.

Let $\mathbf{x} = [\mathbf{p}_c^T, \mathbf{v}_c^T, \mathbf{c}^T]^T \in \mathbb{R}^{15}$ denote the state vector, where $\mathbf{p}_c$ is the MD's centroid position, $\mathbf{v}_c$ its velocity, and $\mathbf{c}$ the vectorized orientation matrix $\mathbf{C}_b^n$. Given the nonlinearity of the state equations, we evaluate the performance of EKF for state estimation. The state transition model is:

$$\mathbf{x}(k+1) = \mathbf{F}(\mathbf{a}_b, \mathbf{\Omega}_{nb}^b)\mathbf{x}(k),$$

where $\mathbf{F}(\mathbf{a}_b, \mathbf{\Omega}_{nb}^b)$ is:

$$\mathbf{F} = \begin{bmatrix} \mathbf{I}_{3\times3} & \Delta t\mathbf{I}_{3\times3} & \frac{\Delta t^2}{2}\mathbf{a}_b^\times \\ \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} & \Delta t\mathbf{a}_b^\times + \frac{\Delta t^2}{2}(\mathbf{a}_b^\times)(\mathbf{\Omega}_{nb}^{b\ \times}) \\ \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{I}_{9\times9} + \Delta t(\mathbf{\Omega}_{nb}^{b\ \times}) + \frac{\Delta t^2}{2}(\mathbf{\Omega}_{nb}^{2\ \times}) \end{bmatrix}.$$

Here, $\mathbf{a}_b^\times = \mathbf{a}_b^T \otimes \mathbf{I}_{3\times3}$, $(\mathbf{\Omega}_{nb}^b)^\times = (\mathbf{\Omega}_{nb}^b)^T \otimes \mathbf{I}_{3\times3}$, and $(\mathbf{\Omega}_{nb}^{2\ \times}) = (\mathbf{\Omega}_{nb}^b \bigotimes \mathbf{I}_{3\times3})(\mathbf{\Omega}_{nb}^b \bigotimes \mathbf{I}_{3\times3})$, where $\otimes$ denotes Kronecker product operation, and $\mathbf{I}_{3\times3}$ is the identity matrix. The acceleration $\mathbf{a}_b$ is assumed to be free of gravity.

The nonlinearity in the state-transition model arises from the multiplication of the inputs by the state vector. The input noise is modeled as additive Gaussian noise:

$$\mathbf{u}(k) = \begin{bmatrix} \mathbf{a}_b(k) \\ \omega_b(k) \end{bmatrix} = \tilde{\mathbf{u}}(k) + \mathbf{n}(k), \quad (9)$$

where $\tilde{\mathbf{u}}(k)$ is the noiseless input, and $\mathbf{n}(k)$ is zero-mean Gaussian noise with covariance $\mathbf{Q} = \text{diag}(\sigma_{\mathbf{a}_x}^2, \sigma_{\mathbf{a}_y}^2, \sigma_{\mathbf{a}_z}^2, \sigma_{\omega_x}^2, \sigma_{\omega_y}^2, \sigma_{\omega_z}^2)$.

The measurement vector $\mathbf{q} = [\mathbf{p}_1^T, \mathbf{p}_2^T, \mathbf{p}_3^T]^T \in \mathbb{R}^9$ represents the vertices' positions estimated via Riemannian localization. The measurement equation is:

$$\mathbf{q}(k+1) = \mathbf{H}\mathbf{x}(k+1) + \nu(k+1), \quad (10)$$

where the measurement (observation) matrix $\mathbf{H}$ is:

$$\mathbf{H} = \begin{bmatrix} \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{d}_1^T \bigotimes \mathbf{I}_{3\times3} \\ \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{d}_2^T \bigotimes \mathbf{I}_{3\times3} \\ \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{d}_3^T \bigotimes \mathbf{I}_{3\times3} \end{bmatrix}, \quad (11)$$

$\mathbf{d}_i$ represents the vector from the centroid of the triangle to the i^{th} vertex, and $\nu(k)$ is zero-mean Gaussian noise with covariance $\mathbf{R}_k = \text{diag}(\sigma_{\mathbf{p}_{1_x}}^2, \sigma_{\mathbf{p}_{1_y}}^2, \ldots, \sigma_{\mathbf{p}_{3_z}}^2)$. Due to the nonlinear process equation, the linear Kalman filter is suboptimal. This paper evaluates the EKF for estimating the MD's position and orientation.

A. Extended Kalman Filter

The EKF operates in two steps: prediction and correction. Given an initial state estimate $\mathbf{x}(0)$, the state and its covariance matrix are predicted as:

$$\bar{\mathbf{x}}(k+1) = \mathbf{F}(\mathbf{a}_b, \omega_b)\hat{\mathbf{x}}(k) \quad (12)$$

$$\bar{\mathbf{P}}(k+1) = \mathbf{F}(\mathbf{a}_b, \omega_b)\hat{\mathbf{P}}(k)\mathbf{F}^T(\mathbf{a}_b, \omega_b) + \mathbf{F}_\mathbf{u}\mathbf{Q}_k\mathbf{F}_\mathbf{u}^T, \quad (13)$$

where $\bar{\mathbf{x}}(k+1)$ and $\bar{\mathbf{P}}(k+1)$ are the predicted state and covariance at time $k+1$, and $\hat{\mathbf{x}}(k)$ and $\hat{\mathbf{P}}(k)$ are the corrected state and covariance at time k. The Jacobian matrix $\mathbf{F}_\mathbf{u}$ of $\mathbf{F}(\mathbf{a}_b, \omega_b)$ with respect to $\mathbf{p}_c$, $\mathbf{v}_c$, and $\mathbf{c}$ is:

$$\mathbf{F}_\mathbf{u} = \begin{bmatrix} \nabla_{\mathbf{a}_b}\mathbf{p}_c^T & \nabla_{\omega_b}\mathbf{p}_c^T \\ \nabla_{\mathbf{a}_b}\mathbf{v}_c^T & \nabla_{\omega_b}\mathbf{v}_c^T \\ \nabla_{\mathbf{a}_b}\mathbf{c}^T & \nabla_{\omega_b}\mathbf{c}^T \end{bmatrix}, \quad (14)$$

with elements:

$$\nabla_{\mathbf{a}_b}\mathbf{p}_c^T = \frac{\Delta t^2}{2}\mathbf{C}_b^n, \qquad \nabla_{\omega_b}\mathbf{p}_c^T = \mathbf{0}_{3\times3}, \qquad \nabla_{\mathbf{a}_b}\mathbf{c}^T = \mathbf{0}_{9\times3},$$

$$\nabla_{\omega_b}\mathbf{c}^T = \Delta t[\mathbf{E}_1^{\times^T}\mathbf{c},\ \mathbf{E}_2^{\times^T}\mathbf{c},\ \mathbf{E}_3^{\times^T}\mathbf{c}] +$$
$$\frac{\Delta t^2}{2}[(_x\mathbf{\Omega}_{nb}^{2\ \times})\mathbf{c},\ (_y\mathbf{\Omega}_{nb}^{2\ \times})\mathbf{c},\ (_z\mathbf{\Omega}_{nb}^{2\ \times})\mathbf{c}],$$

$$\nabla_{\mathbf{a}_b}\mathbf{v}_c = \Delta t\mathbf{C}_b^n + \frac{\Delta t^2}{2}\mathbf{C}_b^n\mathbf{\Omega}_{nb}^b, \quad \nabla_{\omega_b}\mathbf{v}_c = \frac{\Delta t^2}{2}\mathbf{C}_b^n\mathbf{\Xi}_b^T$$

Here, $\mathbf{A}^\times$ denotes the Kronecker product of $\mathbf{A}$ with $\mathbf{I}_{3\times3}$, $\mathbf{e}_i$ are basis vectors, $\mathbf{E}_i$ are their skew-symmetric matrices, and $\mathbf{\Xi}_b^T$ is:

$$\mathbf{\Xi}_b^T = \begin{bmatrix} 0 & \mathbf{a}_{bz} & -\mathbf{a}_{by} \\ -\mathbf{a}_{bz} & 0 & \mathbf{a}_{bx} \\ \mathbf{a}_{by} & -\mathbf{a}_{bx} & 0 \end{bmatrix}.$$

The matrices $_x\mathbf{\Omega}_{nb}^2$, $_y\mathbf{\Omega}_{nb}^2$, and $_z\mathbf{\Omega}_{nb}^2$ are the partial derivatives:

$$_x\mathbf{\Omega}_{nb}^2 = \frac{\partial}{\partial\omega_x}(\mathbf{\Omega}_{nb}^2), \quad _y\mathbf{\Omega}_{nb}^2 = \frac{\partial}{\partial\omega_y}(\mathbf{\Omega}_{nb}^2), \quad _z\mathbf{\Omega}_{nb}^2 = \frac{\partial}{\partial\omega_z}(\mathbf{\Omega}_{nb}^2).$$

To correct the predicted state and covariance, compute the Kalman gain $\mathbf{K}_{k+1}$:

$$\mathbf{K}_{k+1} = \bar{\mathbf{P}}(k+1)\mathbf{H}^T(\mathbf{H}\bar{\mathbf{P}}(k+1)\mathbf{H}^T + \mathbf{R}_{k+1})^{-1} \quad (15)$$

Finally, update the state and covariance:

$$\hat{\mathbf{x}}(k+1) = \bar{\mathbf{x}}(k+1) + \mathbf{K}_{k+1}(\mathbf{q}(k+1) - \mathbf{H}\bar{\mathbf{x}}(k+1))$$
$$\hat{\mathbf{P}}(k+1) = (\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H})\bar{\mathbf{P}}(k+1).$$

B. Projection of the Estimated State to the Manifold

The estimated position and orientation matrix are transformed into the positions of the three vertices of a triangle. However, the EKF does not guarantee that these vertices lie on the isosceles triangle manifold. To address this, a novel and efficient retraction algorithm is introduced to project the estimated vertices onto the manifold.

While the retraction operator in [5] achieves high localization accuracy, it assigns disproportionate weight to one vertex when the term $\mathbf{p}_1^T(\mathbf{p}_2 - \mathbf{p}_3)$ approaches zero, which is inefficient near the true solution. This paper proposes a more efficient retraction method for points close to the optimal

solution. Given high sampling rates, the EKF output, though not on the manifold, is assumed to be near the optimal solution. Thus, the retraction minimally perturbs the output to project it onto the manifold.

Let $\mathbf{Z}^k = [\mathbf{z}_1^k, \mathbf{z}_2^k, \mathbf{z}_3^k]$ denote the vertices obtained from the EKF. The retraction algorithm consists of four steps:

1. **Centering**: Translate the centroid of $\mathbf{Z}^k$ to the origin:

$$\mathbf{z}_i' = \mathbf{z}_i^k - \mathbf{z}_c, \quad \text{where} \quad \mathbf{z}_c = \frac{\mathbf{z}_1^k + \mathbf{z}_2^k + \mathbf{z}_3^k}{3}.$$

2. **Scaling**: Scale the triangle so the base length is d:

$$\mathbf{z}_i'' = \lambda^k \mathbf{z}_i', \quad \text{where} \quad \lambda^k = \frac{d}{||\mathbf{z}_2' - \mathbf{z}_3'||}.$$

3. **Alignment**: Translate $\mathbf{z}_1''$ in the direction of the unit vector $\mathbf{u}_z = (\mathbf{z}_3'' - \mathbf{z}_2'')/d$ such that the translated point $\mathbf{z}_1'''$ falls onto the bisector of the triangle's base. The result of this operation is an isosceles triangle centered around the origin. The median $[\mathbf{z}_1''\mathbf{z}_{2,3}'']$, from the vertex $\mathbf{z}_1''$ that bisects the base is $\mathbf{z}_1'' - (\mathbf{z}_3'' + \mathbf{z}_2'')/2$. The magnitude of the translation γ^k is simply found as the projection of the median $[\mathbf{z}_1''\mathbf{z}_{2,3}'']$ onto the unit vector $\mathbf{u}_z$:

$$\gamma^k = \left\langle \mathbf{z}_1'' - \frac{\mathbf{z}_3'' + \mathbf{z}_2''}{2}, \mathbf{u}_z \right\rangle,$$

where $\langle \cdot, \cdot \rangle$ is the inner product of vectors. The updated positions of the vertices after this step are:

$$\mathbf{z}_1''' = \mathbf{z}_1'' + \gamma^k \mathbf{u}_z, \quad \mathbf{z}_2''' = \mathbf{z}_2'', \quad \mathbf{z}_3''' = \mathbf{z}_3''.$$

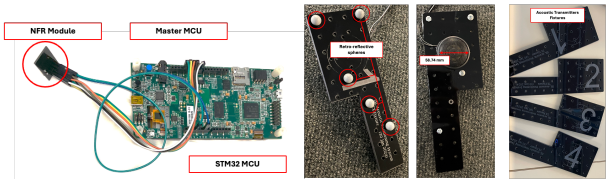
4. **Recentering**: Translate the vertices back by $\mathbf{z}_c$:

$$\mathbf{p}_i^R = \mathbf{z}_i''' + \mathbf{z}_c.$$

The final output is an isosceles triangle with base length d.

V. Experimental Setup

The acoustic localization system consists of three main components: the Master Station, Acoustic Receiver, and Acoustic Transmitters. The Master Station (Figure 2a) uses an STM32F469 Discovery kit with an NRF24L01 module. Each Acoustic Transmitter (Figure 2b) also employs an STM32F469 Discovery kit, an XM-N1004 Sony stereo amplifier, and two Pioneer TS-T110 tweeters (7 kHz bandwidth, 20 kHz central frequency) mounted on tripods with retro-reflective spheres.



(a) Master Station (b) Acoustic transmitter

Fig. 2: Master and Acoustic Transmitter

The Acoustic Receiver (Figure 3), representing the MD, includes an STM32F469 Discovery kit, an NRF24L01 module, an MTi-1 IMU kit, and a custom PCB with four microphones arranged at the vertices of two equilateral triangles (side length: 36.80 mm). Only three microphones are used in this study.

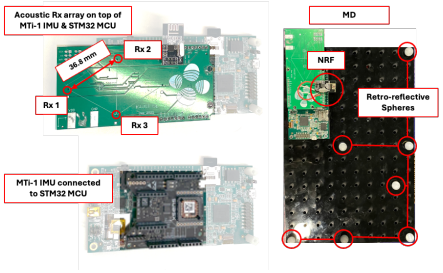


Fig. 3: Acoustic receiver

The experimental setup (Figure 4) includes the Master Station, two Acoustic Transmitter Stations, the MD, and an OptiTrack motion capture system [8] for ground truth position and orientation tracking. Retro-reflective markers on the Transmitters and Receiver ensure accurate tracking.

The Master Station establishes a common time reference via a wireless link, triggering acoustic transmissions. The Receiver records signals from three microphones and IMU data. Post-experiment, OptiTrack records are synchronized with Receiver data for error calculation and analysis.

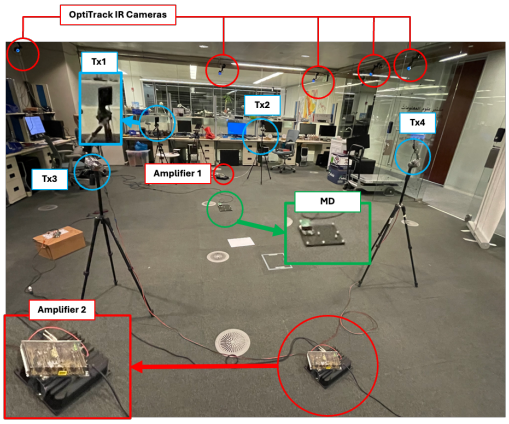
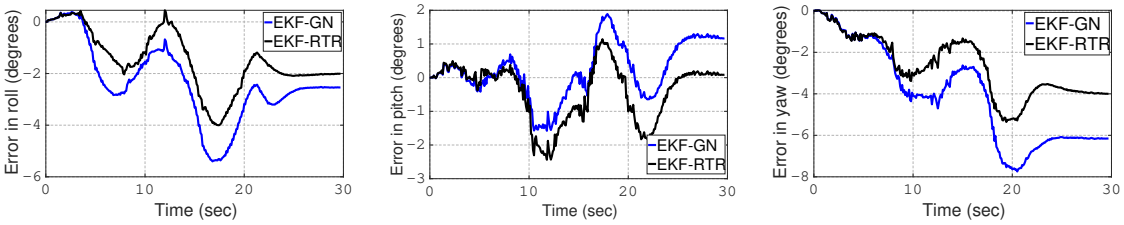


Fig. 4: Experimental setup

VI. Experimental Results

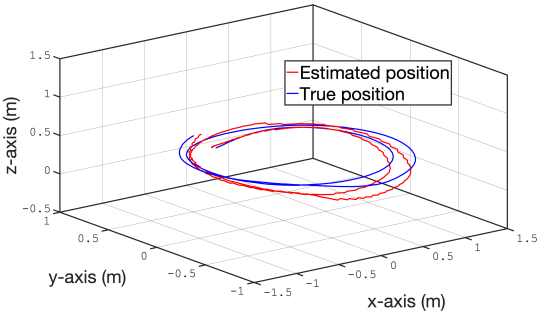
In this experiment, four acoustic transmitters (beacons) were placed at known locations, while three acoustic receivers, arranged in an equilateral triangle on a PCB, were fixed to an MD (Figure 4). The MD was equipped with an MTi-1 IMU, including a triaxial accelerometer and gyroscope. The MD was moved within a 4m × 6m × 3m room (Figure 6). The IMU sampled at 100 Hz, and the acoustic system measured distances between transmitters and receivers at 5 Hz. Synchronization was achieved using an RF signal. The 30-second experiment was recorded and processed offline.

Two Kalman filter variants were evaluated: 1. A Linear Kalman Filter (LKF) with state variables for the MD's centroid position and velocity, using Gauss-Newton (GN) and Riemannian Trust-Region (RTR) methods for vertex position



(a) Error in yaw angle (b) Error in pitch angle (c) Error in roll angle

Fig. 5: Error in the estimated Euler angles.



(a) 3D motion path

Fig. 6: Experimental evaluation: Motion trajectory.

estimation. 2. An EKF with state variables for centroid position, velocity, and orientation matrix, using vertex positions to estimate both orientation and centroid position.

The RMSE of the estimated positions, averaged over the three vertices and time, is shown in Table I. The Riemannian-based EKF outperformed GN-based methods, with over 80% of positions having errors below 9 cm, compared to 13 cm for GN-based EKF as shown in Figure 7.

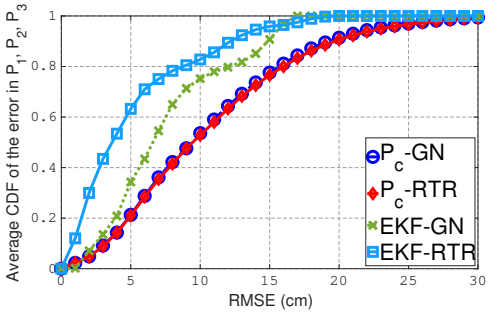


Fig. 7: Error CDF at $\sigma_a = 55$ cm/s^2, $\sigma_\omega = 0.08$ rad/s, $\sigma_d = 12$ cm (estimated from actual measurements.)

TABLE I: RMSE in cm averaged over the three receivers.

Algorithm	LKF-GN	LKF-RTR	EKF-GN	EKF-RTR
RMSE (cm)	10.44	10.60	7.52	5.09

The RMSE of the estimated Euler angles is detailed in Table II. Riemannian-based EKF showed slight improvements over GN-based methods, as illustrated in Figure 5.

TABLE II: RMSE in degrees of the estimated Euler angles.

Algorithm	Yaw RMSE	Pitch RMSE	Roll RMSE
EKF-GN	4.07	0.67	2.40
EKF-RTR	2.68	0.65	1.55

VII. Conclusion

This paper presented an enhanced EKF for high-precision 3D pose estimation of a moving target. The proposed approach integrates measurements from an INS and an acoustic navigation system, with a novel retraction algorithm projecting the filter output onto an isosceles triangle manifold. Experimental results show that using an array of acoustic receivers significantly improves accuracy over single-receiver systems. Additionally, correcting INS orientation with the receiver array yields substantial accuracy gains compared to benchmark methods. Riemannian-based methods, leveraging the receiver array geometry as constraints, outperform Gauss-Newton solvers, as validated through real-world experiments.
Acknowledgments: This research was supported by KAUST and the Ibn Rushd Postdoctoral Fellowship.

References

[1] Y. Sun, W. Wang, L. Mottola, J. Zhang, R. Wang, and Y. He, "Indoor drone localization and tracking based on acoustic inertial measurement," *IEEE Transactions on Mobile Computing*, 2023.

[2] D. Feng, C. Wang, C. He, Y. Zhuang, and X.-G. Xia, "Kalman-filter-based integration of imu and uwb for high-accuracy indoor positioning and navigation," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3133–3146, 2020.

[3] F. Khelifi, A. Bradai, A. Benslimane, P. Rawat, and M. Atri, "A survey of localization systems in internet of things," *Mobile Networks and Applications*, vol. 24, no. 3, pp. 761–785, 2019.

[4] D. Titterton, J. L. Weston, and J. Weston, *Strapdown inertial navigation technology*. IET, 2004, vol. 17.

[5] M. H. AlSharif, A. Douik, M. Ahmed, T. Y. Al-Naffouri, and B. Hassibi, "Manifold optimization for high-accuracy spatial location estimation using ultrasound waves," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5078–5093, 2021.

[6] M. H. AlSharif, M. Saad, M. Siala, T. Ballal, H. Boujemaa, and T. Y. Al-Naffouri, "Zadoff-Chu coded ultrasonic signal for accurate range estimation," in *Signal Processing Conference (EUSIPCO), 2017 25th European*. IEEE, 2017, pp. 1250–1254.

[7] M. H. AlSharif, M. Saad, M. Siala, M. Ahmed, and T. Y. Al-Naffouri, "Range estimation of a moving target using ultrasound differential Zadoff-Chu codes," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–15, 2021.

[8] https://optitrack.com/cameras/primex 41/. (2021) Optitrack system overview. [Online]. Available: https://optitrack.com/cameras/primex-41/