Pose Graph Optimization for a MAV Indoor Localization Fusing 5GNR TOA with an IMU

Meisam Kabiri*, Claudio Cimarelli*, Hriday Bavle*, Jose Luis Sanchez-Lopez*, and Holger Voos*†

* *Interdisciplinary Center for Security Reliability and Trust (SnT)*
University of Luxembourg, Luxembourg
† *Faculty of Science, Technology, and Medicine (FSTM), Department of Engineering*
University of Luxembourg, Luxembourg
Email: {meisam.kabiri, claudio.cimarelli, hriday.bavle, joseluis.sanchezlopez, holger.voos}@uni.lu

Abstract—This paper explores the potential of 5G new radio (NR) Time-of-Arrival (TOA) data for indoor drone localization under different scenarios and conditions when fused with inertial measurement unit (IMU) data. Our approach involves performing graph-based optimization to estimate the drone's position and orientation from the multiple sensor measurements. Due to the lack of real-world data, we use Matlab 5G toolbox and QuaDRiGa (quasi-deterministic radio channel generator) channel simulator to generate TOA measurements for the EuRoC MAV indoor dataset that provides IMU readings and ground truths 6DoF poses of a flying drone. Hence, we create twelve sequences combining three predefined indoor scenarios setups of QuaDRiGa with 2 to 5 base station antennas. Therefore, experimental results demonstrate that, for a sufficient number of base stations and a high bandwidth 5G configuration, the pose graph optimization approach achieves accurate drone localization, with an average error of less than 15 cm on the overall trajectory. Furthermore, the adopted graph-based optimization algorithm is fast and can be easily implemented for onboard real-time pose tracking on a micro aerial vehicle (MAV).

Index Terms—5G TOA, IMU, QuaDRiGa, Indoor Localization, Pose Graph Optimization, Sensor Fusion, Micro Aerial Vehicles.

I. INTRODUCTION

Drones, such as micro aerial vehicles (MAVs), have become increasingly prevalent in indoor environments due to their potential in various applications, from surveillance to delivery services. For many of these applications, accurate drone positioning and orientation are essential.

Global Navigation Satellite Systems (GNSS), the most widely used positioning technology, encounters challenges in penetrating indoor environments due to signal attenuation and multipath effects. Inertial navigation systems (INSs) are another widely used method for indoor localization, but they can accumulate noise over time, resulting in significant position errors if left uncorrected.

Alternative indoor localization techniques, including Wi-Fi, Bluetooth, and Ultra-Wideband (UWB), exhibit limitations in accuracy, scalability, energy efficiency, and cost [1, 2]. For instance, Wi-Fi is highly susceptible to noise, Bluetooth has restricted range and accuracy, and UWB faces slow progress in standard development. Moreover, Zigbee's emphasis on low-power communication and its limited range further constrain its localization capabilities. Consequently, a need for alternative indoor positioning technologies arises that can offer high accuracy and reliability without relying on GNSS.

Recent advances in wireless communication technologies have paved the way for developing location-based services and applications that rely on accurate localization. In particular, deploying fifth-generation (5G) cellular networks has opened up new possibilities for indoor localization due to their high bandwidth, low latency, and improved coverage [3], with small cell technologies, like femtocells and picocells, facilitating indoor coverage. For downlink positioning, 5G utilizes a dedicated pilot signal called Positioning Reference Signal (PRS), which measures signal delay by correlating the received PRS with a locally generated PRS. The delay, also called Time-of-Arrival (TOA), is then calculated by identifying the peak correlation value between the two signals.

Accurate and reliable indoor localization remains challenging due to the complex and dynamic nature of indoor environments. In this context, our work proposes a novel approach for indoor localization using 5G TOA measurements. However, 5G TOA alone may not provide sufficient information for reliable indoor localization. Thus, our work aims to fuse 5G TOA with an inertial measurement unit (IMU), to improve the real-time pose estimation of a flying MAV. IMUs can measure angular velocities, and linear acceleration, providing complementary information to TOA measurements. To accomplish this, we extract distances from the PRS correlation profiles and use them to formulate a range error function tightly integrated with the IMU measurements in a graph-based optimization technique. Due to the lack of real 5G data recorded from MAVs, we utilize the QuaDRiGa simulator [4] to create accurate 5G TOA measurements based on ground truth 6 Degrees of Freedom (DoF) pose data to ensure precise channel modeling. Hence, our experimental simulations include three indoor configurations with LOS and varying base stations (BSs) numbers virtually placed inside the Euroc MAV dataset environment [5].

To summarize, the contributions provided by this paper are the following:

- Formulation of a factor graph model that tightly optimizes 5G TOA ranges with IMU measurements.

- Evaluation in a state-of-the-art dataset shows accuracy close to centimeter precision while running an efficient real-time algorithm.
- Simulation of 5G TOA comparing multiple antenna and communication settings in an indoor scenario to find the best configuration for precise localization.

II. RELATED WORKS

The literature on localization using 5G is relatively limited, especially considering the mechanical aspects and sensor fusion framework. The existing literature often relies on simplistic methodologies and scenarios. Ferre et al. [6] compared localization accuracy for different combinations of the 5G network configurations (center frequency, sub-carrier spacing, and PRS comb size) in terms of the Root Mean Square Error (RMSE). Also, their study considered a fixed target and employed multilateration based on PRS-derived TOA from multiple BSs. A study by del Peral-Rosado et al. [7] explored the impact of positioning performance when BSs are linearly placed along a straight roadside 5G network. They utilized Gauss-Newton optimization and simulated a 100 km/h vehicle on a highway. The study revealed an accuracy of less than 20-25 cm for a communication bandwidth of 50-100 MHz. Additionally, the researchers calculated the TOA by determining the first correlation peak between the PRS and the received signal. Saleh et al. [8] proposed a time-based positioning by combining vehicle velocity information and 5G measurements. They evaluated their approach in a simulated urban canyon using Siradel's S_5GChannel simulator and employed an EKF with a constant velocity model for sensor fusion. The study also analyzed the impact of the 5G geometrical setup on EKF positioning estimation. Another EKF-based positioning framework is proposed by Menta et al. [9]. The authors leveraged the 5G Angle of Arrival (AOA) extracted from the communication signal of BSs equipped with multi-array antennas. By utilizing this information, they achieved sub-meter accuracy in localization. Sun et al. [10] studied localization by combining AOA estimates from 5G BSs with TOA measurements from GNSS satellites. The authors utilized the Taylor series to linearize the mathematical model. As post-processing, they applied a moving averaging to the raw position estimates to minimize errors. Finally, [11] explored the fusion of beamformed RSS information with GNSS data using Neural Networks (NN), achieving meter accuracy.

Unlike previous approaches, we address indoor localization using the factor graph to model the relation among non-homogenous sensor measurements. Furthermore, we leverage the advanced IMU pre-integration factor to propagate the MAV's 6 DoF pose between two lower frequency TOA measurements, obtaining 6DoF pose estimates at a higher frequency.

III. METHODOLOGY

This section describes the proposed drone localization approach in 5G networks. This involves the fusion of range measurements, which we obtain from the TOA of the signal from multiple BSs, with IMU data, *i.e.,* angular velocity and linear acceleration measurements.

A. 5G ToA Estimation

To estimate the distance to 5G BSs, we must first obtain the ToA values along the robot's trajectory. Since we do not possess actual data on 5G communication, this step first requires generating the 5G signal transmitted by each base station, including PRS symbols. Then, a channel simulator creates an impulse response that emulates the wireless channel characteristics based on specific network configurations and given receiver and transmitter positions. Convolving the transmitted signal with the impulse response replicates the effects of the transmission environment, generating the received signal.

At the receiver, the signal is correlated with the corresponding PRS of each base station resulting in a PRS correlation profile. By analyzing it, we identify the TOA as one of the peak or local maxima. Notably, the real TOA is a value close to the best peak but often not matching it. We apply a heuristic selection of the first peak surpassing a global threshold that we find experimentally on the data.

B. Pose Estimation

To accurately estimate the drone's 6DoF pose, we use a graph-based optimization technique that models the relationships between the pose variables using sensor measurements and optimizes the estimation using non-linear least squares. This approach involves creating a factor graph model [12] (see the one abstracting our problem formulation in Figure 1) where the nodes represent the state variables to be estimated and the edges, called factors, represent residual error functions that compare predicted states and observed measurements. Therefore, the factor graph models the posterior probability density $p(\mathcal{X}|\mathcal{Z})$ of the state variables $\mathcal{X}$ given a set of measurements $\mathcal{Z}$, and, assuming independent measurements, it can be factorized into likelihoods $p(\mathcal{Z}_t^f|\mathcal{X}_t)$ and prior $p(\mathcal{X}_0)$:

$$p(\mathcal{X}|\mathcal{Z}) \propto p(\mathcal{X}_0)p(\mathcal{Z}|\mathcal{X}) = p(\mathcal{X}_0) \prod_{\forall t \in \mathcal{T}, \forall f \in \mathcal{F}} p(\mathcal{Z}_t^f|\mathcal{X}_t), \quad (1)$$

where $\mathcal{F}$ defines the set of factors type of functions that can replace the likelihoods. We denote the set of measurements the factor f uses for computing the residual at time t with $\mathcal{Z}_t^f$, where $\mathcal{T}$ is the set of tracked time frames.

1) State Variables: We aim to determine the 3D location and orientation of the MAV's body center, which we align with the IMU frame. As we track the full history of 6DoF poses, with a full-smoothing approach, the set of state variables $\mathcal{X}$ contains all the poses from the start of the trajectory T_1 to the end T_N, where N is the total number of pose nodes added to the graph. Especially, each $\mathsf{T}_t \doteq (\mathsf{R}, \mathbf{p}) \in \mathrm{SE}(3), \forall t \in \mathcal{T} = \{1, \ldots, N\}$ is composed of a rotation $\mathsf{R}_t \in \mathrm{SO}(3)$ and a translation $\mathbf{p}_t \in \mathbb{R}^3$ that transforms the body frame B to the world frame W, where the BSs are placed, at time t.

Additional to the 6DoF transformations, $\mathcal{X}$ comprises the MAV's linear velocity $\mathbf{v}_t \in \mathbb{R}^3$. We also need to estimate the time-variant biases of the IMU's gyroscope $\mathbf{b}_t^g \in \mathbb{R}^3$ and
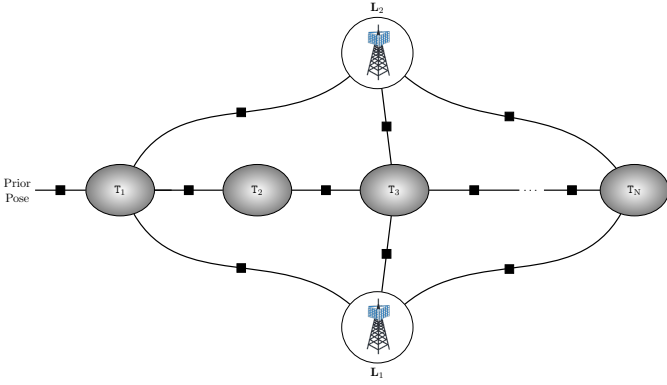
Fig. 1. The figure visualizes the structure of the factor graph used to optimize the variables, represented by circles, by relating them through factors, represented by squares. The nodes $\mathtt{T}_t$ contain the 6DoF pose variables connected by IMU pre-integration factors (the bias and velocity nodes are not visualized). TOA measurements create range factors between robot pose nodes and BSs position nodes, $\mathbf{L}_1$ and $\mathbf{L}_2$. A prior factor is connected to the first node $\mathtt{T}_1$ to constrain it with the initial trajectory pose.

accelerometer $\mathbf{b}_t^a \in \mathbb{R}^3$ to keep track of the IMU noise drift. Lastly, we include the 3D positions of the 5G antennas $\mathbf{L}_k \in \mathbb{R}^3, \forall k \in \{1, \dots, \mathrm{K}\}$, where K is the total number of BSs.

2) IMU Factor: Our approach involves a 6-axis IMU that measures the body B linear acceleration ${}_\mathtt{B}\tilde{\mathbf{a}}_t$ and angular velocity ${}_\mathtt{B}\tilde{\boldsymbol{\omega}}_t$ expressed in the W frame. The IMU real motion state $\{{}_\mathtt{B}\mathbf{a}_t, {}_\mathtt{B}\boldsymbol{\omega}_t\}$ is altered by additive Gaussian white noise $\{\boldsymbol{\eta}_t^a, \boldsymbol{\eta}_t^g\}$ and slowly time-varying biases $\{\mathbf{b}_t^a, \mathbf{b}_t^g\}$ affecting respectively the accelerometer and gyroscope as defined by the following IMU model:

$$\tag{2} {}_\mathtt{B}\tilde{\boldsymbol{\omega}}_t = {}_\mathtt{B}\boldsymbol{\omega}_t + \mathbf{b}_t^g + \boldsymbol{\eta}_t^g$$

$$\tag{3} {}_\mathtt{B}\tilde{\mathbf{a}}_t = {}_\mathtt{B}\mathbf{a}_t - \mathtt{R}_t^\mathsf{T}\mathbf{g} + \mathbf{b}_t^a + \boldsymbol{\eta}_t^a \,,$$

where $\mathbf{g}$ is the Earth's gravity vector in the world frame W.

Due to the IMU's higher sampling frequency than other sensors, it typically captures multiple measurements between two TOA instances. The IMU factor is constructed utilizing a *preintegrated measurement* [13] constraining the relative motion increments. Especially, we obtain the condensed measurements $\Delta\tilde{\mathtt{R}}_{ij}$ of rotation, $\Delta\tilde{\mathbf{p}}_{ij}$ of position, and $\Delta\tilde{\mathbf{v}}_{ij}$ of velocity by integrating multiple IMU readings $\{{}_\mathtt{B}\tilde{\mathbf{a}}_t, {}_\mathtt{B}\tilde{\boldsymbol{\omega}}_t : \forall t \in [i, j]\}$. So, we can define the residual terms $\mathbf{r}$ for the states $\{\mathtt{R}_{ij}, \mathbf{p}_{ij}, \mathbf{v}_{ij}\}$:

$$\tag{4} \mathbf{r}_{ij}^\mathtt{R} \doteq \mathrm{Log}\left(\Delta\tilde{\mathtt{R}}_{ij}^\mathsf{T} \, \mathtt{R}_i^\mathsf{T}\mathtt{R}_j\right) \,,$$

$$\tag{5} \mathbf{r}_{ij}^\mathbf{P} \doteq \mathtt{R}_i^\mathsf{T}\left(\mathbf{p}_j - \mathbf{p}_i - \mathbf{v}_i\Delta t_{ij} - \tfrac{1}{2}\mathbf{g}\Delta t_{ij}^2\right) - \Delta\tilde{\mathbf{p}}_{ij} \,,$$

$$\tag{6} \mathbf{r}_{ij}^\mathbf{v} \doteq \mathtt{R}_i^\mathsf{T}\left(\mathbf{v}_j - \mathbf{v}_i - \mathbf{g}\Delta t_{ij}\right) - \Delta\tilde{\mathbf{v}}_{ij} \,,$$

where Δt_{ij} is the total time interval. Also, $\mathrm{Log} : \mathrm{SO}(3) \to \mathbb{R}^3$ defines the logarithm map that associates elements of the rotation manifold $\mathrm{SO}(3)$ to vectors on the Euclidean tangent space $\mathbb{R}^3$ representing rotation increments.

Regarding the biases, the total residual $\mathbf{r}_{ij}^\mathbf{b}$ between time $t = i$ and $t = j$ and $i < j$ is set as follows:

$$\tag{7} \mathbf{r}_{ij}^\mathbf{b} \doteq \mathbf{b}_j^g - \mathbf{b}_i^g + \mathbf{b}_j^a - \mathbf{b}_i^a \,.$$

3) TOA Range Factor: By multiplying the estimated TOA values δ_{sk} by the speed of light c, *i.e.*, $d_{sk} = \delta_{sk} \cdot c$, we obtain K metric distance measurements $d_{sk} \in \mathbb{R}, \forall s \in \mathcal{S} \subseteq \mathcal{T}$ of the drone to the k-th landmark $\mathbf{L}_k$ at time $s = i$. Notably, we explicitly express the possibility of having fewer TOA measurements than the number of tracked poses. The residual $\mathbf{r}_{ik}^\delta$ of the TOA factor at time $s = i$ with the BS $\mathbf{L}_k$ is defined as:

$$\tag{8} \mathbf{r}_{ik}^\delta \doteq d_{ik} - \|\mathbf{p}_i - \mathbf{L}_k\|_2 \,.$$

4) Optimization: The optimization problem is formulated as Maximum a Posteriori Estimation (MAP) estimation that involves finding the state $\mathcal{X}^*$ that maximizes the posterior:

$$\tag{9} \mathcal{X}^* = \arg\max_\mathcal{X} p(\mathcal{X}|\mathcal{Z}).$$

Considering the proportional relationship in Equation 1, Equation 9 is equivalent to the maximization of the problem factorized through likelihood functions:

$$\tag{10} \mathcal{X}^* = \arg\max_\mathcal{X} p(\mathcal{X}_0) \prod_{\forall t \in \mathcal{T}, \forall f \in \mathcal{F}} p(\mathcal{Z}_t^f|\mathcal{X}_t) \,.$$

Notably, with factor graphs, likelihoods can be expressed by the more general factors, which we have defined with $f \in \mathcal{F} = \{\mathtt{R}, \mathbf{p}, \mathbf{v}, \mathbf{b}, \delta\}$ referring to the related residual functions. By assuming that the measurements' errors are zero-mean Gaussian distributed, Equation 10 is analogous to the minimization of the negative log-likelihood:

$$\tag{11} \mathcal{X}^* = \arg\min_\mathcal{X} \|\mathbf{r}_0\|_{\boldsymbol{\Sigma}_0}^2 + \sum_{t=1}^\mathrm{N} \sum_{\forall f \in \mathcal{F}} \left\|\mathbf{r}_t^f\right\|_{\boldsymbol{\Sigma}_t^f}^2 \,,$$

where $\|\mathbf{r}\|_{\boldsymbol{\Sigma}}^2 = \mathbf{r}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{r}$ is the squared Mahalanobis norm, and $\mathbf{r}_t^f$ are the residual functions of the aforementioned factors f computed at time t with covariance matrix $\boldsymbol{\Sigma}_t^f$. We denote with $\mathbf{r}_0$ the residual derived from the prior on the initial pose with $\boldsymbol{\Sigma}_0$ being its covariance matrix.

To efficiently solve the MAP optimization problem, we utilize the iSAM2 iterative optimization algorithm [14] implemented in GTSAM [15]. This algorithm can automatically identify the variables that require linearization at each step, and it enables us to keep our graph solution updated while adding new nodes without experiencing memory overload.

IV. EXPERIMENTS

A. Simulation of the 5G Communication

This study uses 5G specifications for indoor base stations to generate PRS signals for TOA signals. Our simulation relies on the MATLAB 5G Toolbox to generate the resource grids for 5G NR signals, including the PRS and Physical Downlink Shared Channel (PDSCH) resources.

Aiming at a realistic simulation of the 5G NR wireless communication, we employ QuaDRiGa (quasi-deterministic radio channel generator) channel simulator [4]. For precise channel modeling, it is essential to have information regarding the drone's position, orientation, and velocity. This information is required to account for factors such as the Doppler shift

effect accurately. We also assume that both the receiver and all transmitters use omnidirectional antennas. Three 5G simulation scenarios were considered, with different frequencies: QuaDRiGa-Industrial-LOS for 5 GHz, 3GPP-38.901-Indoor-LOS for 28 GHz, and mmMAGIC-Indoor-LOS for 78 GHz.

- **QuaDRiGa_Industrial_LOS [16]**: This scenario is designed to replicate a LOS environment for industrial applications.
- **3GPP_38.901_Indoor_LOS [17]**: This scenario simulates indoor environments, *e.g.,* office buildings and shopping centers, with 0.5-100 GHz LOS frequency.
- **mmMAGIC_Indoor_LOS [18]**: This is designed specifically for frequencies ranging from 6-100 GHz and indoor environments, *e.g.,* offices, with LOS.

The configurations for each 5G simulation scenario are provided in Table I.

TABLE I
5G SIMULATION SCENARIOS' SYSTEM CONFIGURATIONS.

Scenario	Freq. (GHz)	Bandwidth (MHz)	Subcarrier Spacing (KHz)	Num. of RBs	Comb Size	SNR	Cyclic Prefix
QuaDRiGa-Industrial-LOS	5 (FR1)	100	30	275	6	10 dB	normal
3GPP-38.901-Indoor-LOS	28 (FR2)	200	60	275	6	10 dB	normal
mmMAGIC-Indoor-LOS	78 (FR2)	400	120	275	6	10 dB	normal

B. Experimental Environment

As the 5G channel simulation requires knowing the state of the receiver, *i.e.,* its pose and velocity, we require a flying drone dataset to evaluate our method. The EuRoC MAV dataset [5] is a widely used benchmark dataset for visual-inertial odometry and SLAM. It was collected by an indoor drone equipped with a stereo-camera module providing images at 20 Hz and a calibrated IMU at 200HZ. The EuRoC MAV dataset contained the drone's position and orientation data obtained through the Vicon motion capture system, which can record the full 6DoF at about 100Hz. The full set of calibrated rigid transformations between sensors and the Vicon is also given. EuRoC MAV consists of several sequences. For this study, we consider the Vicon Room 1, sequence 01.

We employ QuaDRiGa to model the wireless communication channel based on the available 6DoF ground truth poses of the drone provided by the dataset from which we compute the required velocity considering the translation vectors between two time-consecutive poses. To this aim, we virtually place two to five fictitious BSs in the room where the trajectory is recorded. The positions of the BSs in the EuRoC MAV Vicon system's coordinate frame are $BS_1 = (-10, -7, 2)$, $BS_2 = (7, 13, 3)$, $BS_3 = (25, -35, 4)$, $BS_4 = (-6, 9, 5)$, $BS_5 = (-4, -14, 6)$. We use these values to initialize the corresponding state variables of the optimization problem with a small covariance.

After generating the resource grids and simulating the channel model, we generated the received signal at the receiver every 0.2 seconds enabling the calculation of TOA with the frequency of 5 Hz. To extract the TOA, the received signal correlated with the transmitter's PRS pattern, and the delay was calculated by analyzing the correlation profile.

Typically, the initial or highest peak is considered as the response. This approach can be compromised by noise. The LOS coefficient may be weaker than the multipath coefficient due to attenuation from non-line-of-sight (NLOS) objects or constructive interference. To address this, a threshold was set to eliminate values below it, and the first peak above the threshold was chosen as the response. It is worth noting that the threshold value was determined through experimentation.
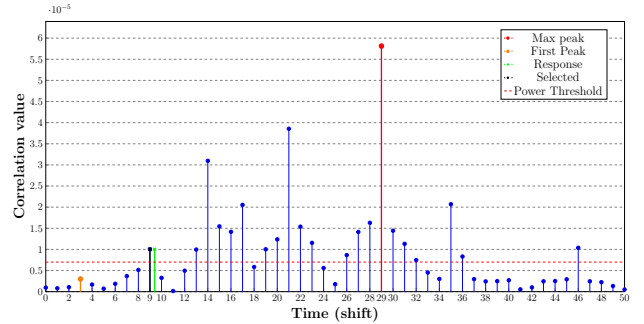

Fig. 2. PRS Correlation Profile

An example of a correlation profile in the simulation is shown in Figure 2, where neither the first nor the maximum peak was the response. Still, a suitable threshold allowed the selection of the first peak as the response. Table II gives the statistic of the error in the resulting estimated distance to each BS.

TABLE II
ERROR STATISTICS OF THE ESTIMATED TOA DISTANCE TO BSs FOR DIFFERENT 5G SIMULATION SCENARIOS GIVEN IN METERS.

5G Sim. Scenario	Statistic	TOA#1	TOA#2	TOA#3	TOA#4	TOA#5
QuaDRiGa-Industrial-LOS	Mean	0.128	-0.044	0.005	-0.080	-0.022
	Std.	0.568	0.810	0.763	0.872	0.717
3GPP-38.901-Indoor-LOS	Mean	-0.024	-0.021	-0.059	0.040	-0.059
	Std.	0.344	0.368	0.351	0.394	0.3690
mmMAGIC-Indoor-LOS	Mean	0. 0.0021	0.0103	-0.007	0.0032	-0.0104
	Std.	0.1845	0.1709	0.1728	0.1592	0.1763

C. Evaluation Metrics

For the evaluation of our approach, we utilize the two most popular metrics in SLAM: Absolute Trajectory Error (ATE) and Relative Pose Error (RPE) of the rotation RPE_R and translation RPE_p [19]. In addition to these metrics, we calculated the RMSE error $E_a, \forall a \in [x, y, z]$ for each trajectory coordinate axis. By calculating the error for each coordinate axis separately, we aim to gain insights into possible differences in accuracy that depend on the spatial direction.

D. Results

The drone's position and orientation results are obtained from the factor graph based on the final MAP estimate for each node. Nodes are generated consistently at 10Hz, twice the TOA's frequency. To evaluate the performance of the localization algorithm, error metrics are computed by comparing the ground truth Vicon pose with the estimated pose that is temporally closest.

5G sim. scenario	BS Num.	ATE(m)	E_x(m)	E_y(m)	E_z(m)	RPE_p(m)	RPE_R(deg)
QuaDRiGa-Industrial-LOS	2	3.7290	0.7326	1.1775	3.4615	0.0174	0.3979
	3	1.2566	0.2978	0.1910	1.2058	0.0113	0.3984
	4	1.5274	0.2230	0.1588	1.5026	0.0091	0.4175
	5	0.6791	0.2431	0.1553	0.6147	0.0082	0.4142
3GPP-38.901-Indoor-LOS	2	2.5805	1.0242	1.6213	1.7266	0.0182	0.3979
	3	1.1189	0.1935	0.1225	1.0952	0.0083	0.3979
	4	0.3787	0.1757	0.1147	0.3152	0.0070	0.4114
	5	0.2583	0.1717	0.0970	0.1668	0.0066	0.4487
mmMAGIC-Indoor-LOS	2	1.9861	0.7814	1.0087	1.5220	0.0104	0.3978
	3	1.2447	0.1228	0.0522	1.2375	0.0065	0.4378
	4	0.1432	0.0576	0.0468	0.1225	0.0055	0.4468
	5	0.1312	0.0516	0.0422	0.1131	0.0052	0.4571

Based on the results, it is evident that increasing the number of BSs enhances the precision of position estimation in all situations except the first simulation scenario when transitioning from three to four. In particular, the mmMAGIC-Indoor-LOS with the highest bandwidth (400 MHz) outperforms the other 5G simulation scenarios.
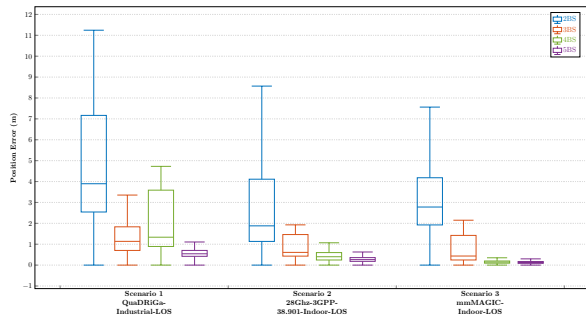
To evaluate the efficiency of the proposed method and its potential for real-time application, we recorded and reported the sum, average, and median optimization times. The sum optimization time was calculated to be 5.203 seconds for the 144-second trajectory of the EuRoC Mav dataset, indicating the total time required to optimize the drone's position and orientation using our graph-based framework. On average, the optimization process took only 0.0036 seconds, demonstrating the method's speed and potential for efficient implementation. Furthermore, the median optimization time was 0.0029 seconds, indicating that most optimization processes were even faster than the average, highlighting the consistency of the algorithm's performance.

All the experiments were performed on a Ubuntu 20.04 laptop with an Intel(R) Core(TM) i9-10885H CPU @ 2.40GHz with 16 cores and 32 Gb of RAM. As the code was partially implemented in Python, we expect further improvement by a complete conversion in C++.
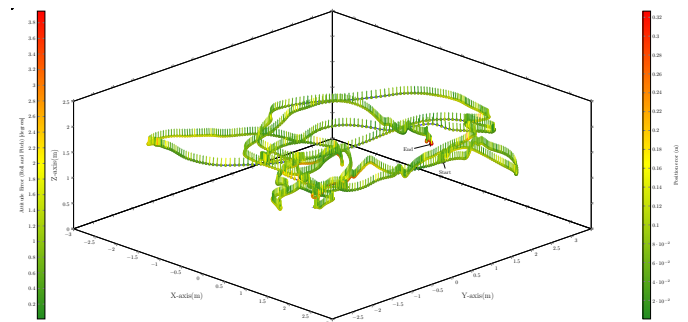

Fig. 3. Box plot of the translation error in meters for each 5G simulation scenario and number of BSs.

Fig. 4. Visualization of the 3D trajectory estimated using five BSs in mmMAGIC-Indoor-LOS 5G simulation scenario. Best viewed online and in color.

Figure 3 illustrates the 3D position error distribution. As expected, the median value of the position error decreases with an increase in the number of BSs. The box plot reveals that the mmMAGIC-Indoor-LOS 5G simulation scenario achieves the lowest overall error. Surprisingly, the first scenario performs slightly worse when transitioning from three to four BSs.

The detailed results are given in Table III. The table includes information on ATE, RPE, and translation RMSE for each motion direction in different 5G simulation scenarios. The error values confirm that a higher number of antennas and a larger bandwidth decrease sensibly the error. However, in the last two best simulation scenarios, the accuracy improvement is marginal by increasing from four to five BSs. This may indicate a lower bound to the error reduction achieved by adding more antennas. Nevertheless, such redundancy may be helpful in those environments where NLOS conditions are more frequent.

In Figure 4, we represent the most accurate results obtained in the 5G simulation scenario mmMAGIC-Indoor-LOS with five BSs. The plot displays the estimated 3D positions of the MAV, with arrows indicating the estimated attitude (excluding yaw). The position and orientation errors are color-coded to show in red the few spots in which the poses do not match well the ground-truth and in green where the error is low, down to a few centimeters.

E. Limitations

The approach used in the study has several limitations and potential for future work. The 3D position is not fully constrained with only two antennas, making convergence difficult without other measurements. Nevertheless, the UAV's rotation errors primarily result from IMU noise, as the radio frequency signal only provides distances to the antennas. The yaw estimation has drift issues because it lacks global measurement to correct it. Integrating other sensors can improve the localization accuracy by observing the rotation around z, *e.g.,* employing a magnetometer. Notably, a camera can be incorporated to add other constraints on the 6DoF relative motion based on visual features and loop closures.

Furthermore, the error in the z axis is larger than along x and y axes because of limited offset or variation in the positions of the base stations in the height direction. We foresee the possibility of fusing the barometer's absolute height measurements to relieve such issues. Additionally, the localization accuracy depends heavily on the quality of the TOA measurements, which can be negatively affected by NLOS conditions. In such cases, correctly setting the measurement uncertainty for each TOA range factor, using Mahalanobis

distance to discard outliers, or applying a robust kernel to the cost function, *e.g.,* Huber, may be beneficial to alleviate the problem.

Finally, the proposed method assumes that the positions of the base stations are known with high confidence and fixed in the exact location, which may not be the case in real-world scenarios where the stations may be moving or their positions may be completely unknown. Moreover, the study assumes that the odometry frame can be initially aligned with the world frame inside which the antennas are placed. This can be solved in future work by explicitly estimating the transformation between the local coordinate frame and the world.

V. Conclusion

In conclusion, we have successfully demonstrated the potential of using 5G TOA-based range measurements with data from inertial sensors to locate a MAV indoors in various scenarios and network setups. Our optimization strategy, which is graph-based, enables us to accurately determine the drone's position and orientation, with an average testing error of less than 15 cm. This technique has many practical applications, such as drone-powered monitoring and communication systems. In the future, we plan to improve localization accuracy and reliability by integrating visual data from cameras, experimenting with real data, and investigating advanced techniques for precise TOA estimation.

References

[1] B. Yang and E. Yang, "A survey on radio frequency based precise localisation technology for uav in gps-denied environment," *Journal of Intelligent & Robotic Systems*, vol. 103, pp. 1–30, 2021.

[2] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[3] M. Kabiri, C. Cimarelli, H. Bavle, J. L. Sanchez-Lopez, and H. Voos, "A review of radio frequency based localisation for aerial and ground robots with 5g future perspectives," *Sensors*, vol. 23, no. 1, p. 188, 2022.

[4] S. Jaeckel, L. Raschkowski, K. Börner, L. Thiele, F. Burkhardt, and E. Eberlein, "Quadriga-quasi deterministic radio channel generator, user manual and documentation," *Fraunhofer Heinrich Hertz Institute, Tech. Rep. v2. 0.0*, 2017.

[5] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, "The euroc micro aerial vehicle datasets," *The International Journal of Robotics Research*, 2016. [Online]. Available: http://ijr.sagepub.com/content/early/2016/01/21/0278364915620033.abstract

[6] R. M. Ferre, G. Seco-Granados, and E. S. Lohan, "Positioning reference signal design for positioning via 5g," *National Committee for Radiology in Finland*, 2019.

[7] J. A. del Peral-Rosado, J. A. López-Salcedo, S. Kim, and G. Seco-Granados, "Feasibility study of 5g-based localization for assisted driving," in *2016 International conference on localization and GNSS (ICL-GNSS)*. IEEE, 2016, pp. 1–6.

[8] S. Saleh, A. S. El-Wakeel, and A. Noureldin, "5g-enabled vehicle positioning using ekf with dynamic covariance matrix tuning," *IEEE Canadian Journal of Electrical and Computer Engineering*, vol. 45, no. 3, pp. 192–198, 2022.

[9] E. Y. Menta, N. Malm, R. Jäntti, K. Ruttik, M. Costa, and K. Leppänen, "On the performance of aoa–based localization in 5g ultra–dense networks," *IEEE Access*, vol. 7, pp. 33 870–33 880, 2019.

[10] C. Sun, H. Zhao, L. Bai, J. W. Cheong, A. G. Dempster, and W. Feng, "Gnss-5g hybrid positioning based on toa/aoa measurements," in *China Satellite Navigation Conference (CSNC) 2020 Proceedings: Volume III*. Springer, 2020, pp. 527–537.

[11] R. Klus, J. Talvitie, and M. Valkama, "Neural network fingerprinting and gnss data fusion for improved localization in 5g," in *2021 International Conference on Localization and GNSS (ICL-GNSS)*. IEEE, 2021, pp. 1–6.

[12] F. Dellaert and M. Kaess, *Factor Graphs for Robot Perception*. Foundations and Trends in Robotics, Vol. 6, 2017. [Online]. Available: http://www.cs.cmu.edu/~kaess/pub/Dellaert17fnt.pdf

[13] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, *IMU preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation*, 2015.

[14] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. Leonard, and F. Dellaert, "isam2: Incremental smoothing and mapping with fluid relinearization and incremental variable reordering," in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 3281–3288.

[15] F. Dellaert and G. Contributors, "borglab/gtsam," May 2022. [Online]. Available: https://github.com/borglab/gtsam)

[16] S. Jaeckel, N. Turay, L. Raschkowski, L. Thiele, R. Vuohtoniemi, M. Sonkki, V. Hovinen, F. Burkhardt, P. Karunakaran, and T. Heyn, "Industrial indoor measurements from 2-6 ghz for the 3gpp-nr and quadriga channel model," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*. IEEE, 2019, pp. 1–7.

[17] 3GPP, "Study on channel model for frequencies from 0.5 to 100 ghz," 2018.

[18] L. Carneiro de Souza, C. H. de Souza Lopes, R. de Cassia Carlleti dos Santos, A. Cerqueira Sodré Junior, and L. L. Mendes, "A study on propagation models for 60 ghz signals in indoor environments," *Frontiers in Communications and Networks*, vol. 2, p. 757842, 2022.

[19] D. Prokhorov, D. Zhukov, O. Barinova, K. Anton, and A. Vorontsova, "Measuring robustness of visual slam," in *2019 16th International Conference on Machine Vision Applications (MVA)*. IEEE, 2019, pp. 1–6.