IEEE Copyright Notice

Visual-inertial SLAM aided estimation of anchor poses and sensor error model parameters of UWB radio modules

Philipp Lutz[1*], Martin J. Schuster[1], Florian Steidle[1]

Abstract— Local positioning technologies based on ultra-wideband (UWB) ranging have become broadly available and accurate enough for various robotic applications. In an infrastructure setup with static anchor radio modules one common problem is to determine their global positions within the world coordinate frame. Furthermore, issues like the complex radio-frequency wave propagation properties make it difficult to design a consistent sensor error model which generalizes well across different anchor setups and environments. Combining radio based local positioning systems with a visual-inertial navigation system (VINS) can provide very accurate pose estimates for calibration of the radio based localization modules and at the same time alleviate the inherent drift in visual-inertial navigation. We propose an approach to utilize a visual-inertial SLAM system using fish-eye stereo cameras and an IMU to estimate the anchor 6D poses as well as the parameters of an UWB module sensor error model on a micro-aerial-vehicle (MAV). Fiducial markers on all anchor radio modules are used as artificial landmarks within the SLAM system to get accurate anchor module pose estimates. *Index Terms*— MAVs, mobile robots, SLAM, UWB, radio localization, sensor calibration"

Fig. 1. The multicopter 'Ardea' equipped with its two pairs of fisheye stereo cameras, an IMU and a UWB radio ranging module with its antenna coordinate frame visualized.

I. INTRODUCTION

Visual-inertial navigation has been employed successfully on mobile robots and specifically on micro-aerial-vehicles (MAVs) to navigate in both indoor and outdoor environments [1], [2]. However, the integration of relative position and orientation estimation from a progressive stream of noisy sensor data leads to an unavoidable drift. State-of-the-art visual-inertial navigation system show a promising absolute trajectory error (ATE) of lower than 0.5 % in typical scenarios [3]. Coupling an VINS with an absolute sensor based on radio localization such as global-navigation-satellite-systems (GNSS) or local positioning systems (LPS) helps to compensate long-range drifts, whereas the VINS provides short-range accuracy.

In this work we use an VINS together with a LPS and show how our SLAM backend helps to calibrate both a UWB sensor measurement error model as well as the static anchor module positions and orientations.

A. UWB radio based localization

The use of local-positioning-systems (LPS) contrary to GNSS for navigation of mobile robots is attractive because they do not have to rely on satellite communication which can be problematic in scenarios where signal reflections and shading can severely corrupt position measurements and hence render them unsuitable for the required application

accuracy. With a LPS one can spread out static beacons (anchor nodes) at arbitrary locations and therefore actively influence the node network geometry to ensure line-of-sight (LOS) conditions and optimize the geometry dependent dilution of precision (DOP) quality metric [4] for the required localization accuracy. There are three common ranging methods: using received signal level (RSL), signal time-of-flight (ToF) and angle-of-arrival (AoA) utilizing antenna arrays. ToF based ranging is a good trade-off between achievable accuracy and equipment complexity, because it requires only one antenna [5]. After all pair-wise ranging measurements between the tag and all anchor nodes are obtained, the set of non-linear equations - the so-called multilateration equations - have to be solved in order to locate the moving robot carrying the tag module. Applying a non-linear least-squares (NLLS) solver to those equations is a straight-forward method to solve the multilateration problem. However, also sophisticated algebraic closed-form solutions for both time-of-arrival [6] as well as hyperbolic time-difference-of-arrival [7] measurements have been studied. A comparison between a weighted iterative NLLS solver and common closed-form solutions is also investigated [6]. Moreover, various parametric and non-parametric bayesian estimators such as Kalman and Particle filters have been successfully used for localization [8]. As for Kalman filters, due to the highly nonlinear and non-gaussian nature of radio wave propagation, the linearization used in EKFs as well as the gaussian error assumption usually lead to higher localization errors compared to particle filter or NLLS approaches [8]. In Section III-C we briefly discuss our NLLS based localization

[1]Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Germany

*Corresponding author: `philipp.lutz@dlr.de`

approach using double-sided two-way-ranging (DS-TWR).

B. UWB radio module measurement error sources and calibration

Theoretically UWB localization has the potential to provide centimeter-level accuracy [9], however in practice sophisticated sensor error models are required to mitigate the effect of biases caused by non-deterministic signal propagation, most notably non-line-of-sight (NLOS) conditions. Typical error sources for UWB radio module based ToF ranging include

- NLOS conditions due to occlusions between tag and anchors,
- poor anchor geometry, which is reflected in the DOP metric [4],
- strongly anisotropic antenna characteristics of small and cheap antennas [10] and their immediate environment,
- antenna delay due to radio wave propagation through circuitry (PCB) and antenna before propagating through air,
- receiver specifics such as ranging biases depending on received signal level.

Reliably detecting NLOS ranging conditions and compensating for NLOS induced errors is very tricky and involves detailed knowledge about the underlying radio module's physical layer and therefore is not considered within this manuscript, only simple outlier checks are performed. One of the other dominant errors is caused by the anisotropic (i.e. directional) properties of small and cheap antennas used on common UWB based ranging modules, for example the used Decawave DWM1000 modules [11]. A model based approach using the LOS signal entrance and exit angles between communicating pairs of ranging module antennas as parameters was presented in [12]. In [10], [13], [14] antenna orientation dependent errors and errors caused by objects in the antenna vicinity are estimated without an explicit model utilizing Gaussian Processes. However assumptions regarding the anchor orientations have been made because they were not observable. Hence we would like to motivate a method for jointly estimating the 6D pose of all anchors as well as the sensor error model parameters for all tag-anchor pairs in one calibration run. The estimation of all anchor antenna poses is done by incorporating our visual-inertial SLAM system with fiducial markers as artificial landmarks on all anchor modules, which is briefly described in Section III. More emphasis is put on considering aforementioned error sources for an sensor error model and the error model which was derived empirically based on our findings, see Section II.

C. Contributions

The contributions of this paper are:

- Introduction to our novel visual navigation pipeline coupled with Decawave DWM1000 UWB radio modules (Section III) on our experimental MAV platform (Section IV).

- UWB sensor measurement error model, focused on errors induced by antenna and immediate antenna environment properties originating from module mount and close objects (Section II).
- Joint calibration of a UWB radio sensor error model and estimating anchor module poses with a SLAM system using fiducial markers (Section II-D).

II. RANGING ERROR MODEL

A. Considered error sources

As introduced in Section I-B we would like to focus on the following sources of error: antenna directivity, antenna delay bias and environment influence around antennas which leads to NLOS or multipath effects. Influences like deviation between tag and anchor clock frequency are not explicitly considered, because they are mostly mitigated already by using a double-sided two-way ranging scheme [15]. Another well understood error source is the temperature depended frequency drift in clock oscillators. The oscillator in the DWM1000 modules is temperature compensated [11], we nevertheless measured tag and anchor module temperatures during several experiments but couldn't find any significant influence on the range measurements when all sensors were exposed to the same ambient temperature. Therefore temperature is not considered in the error model. Small manufacturing and assembly variances can cause deviations in internal propagation delays, resulting in range measurement offsets. These offsets can vary from module to module and are generally identified in a separate calibration routine as outlined in [16]. Antenna related ranging errors are most likely caused by non-isotropic radiation patterns [11] influencing gain and group-delay [17]. Ideal antennas are not directional and therefore exhibit the same properties in all directions. On our experimental hexacopter platform *Ardea* we have components, such as carbon fiber tubes, cameras, PCBs and motors in the immediate vicinity of the UWB antenna, see Fig. 1 and Section IV. This leads to strong directional antenna characteristics for incoming as well as outgoing radio waves at the tag module in case the LOS between the tag and its current communication partner is obscured by robot components. Anchor modules are usually free-standing, without objects in the direct vicinity of their antennas, at least one has more control over the anchor placement compared to mounting the tag module on a robotic platform, far away from other components.

B. Exploration of real data

The setup shown on the left of Fig. 2 is used for exploration of typical ranging errors of the used Decawave DWM1000 UWB modules (see Fig. 8) in a simple experiment where the robot - which carries the tag - is moved randomly in between 4 anchor modules. Anchors as well as the robot are equipped with visual markers and are tracked via a Vicon visual tracking system to provide a ground-truth for calculating ranging errors.

In Fig. 3 we show how the aforementioned errors are distributed and that they do not follow a gaussian distribution,
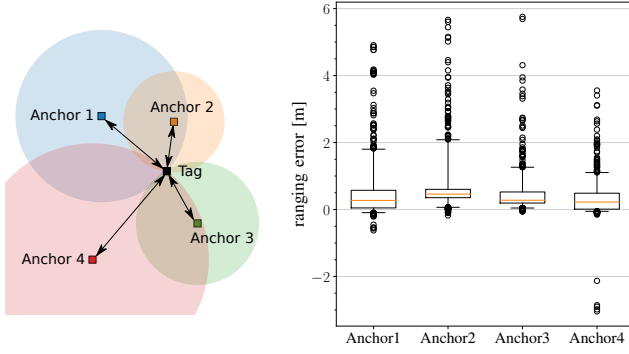
Fig. 2. Ranging test setup with 4 anchors for measurements with a moving tag and the respective ranging errors boxplots with whiskers at percentile 2 and 98.
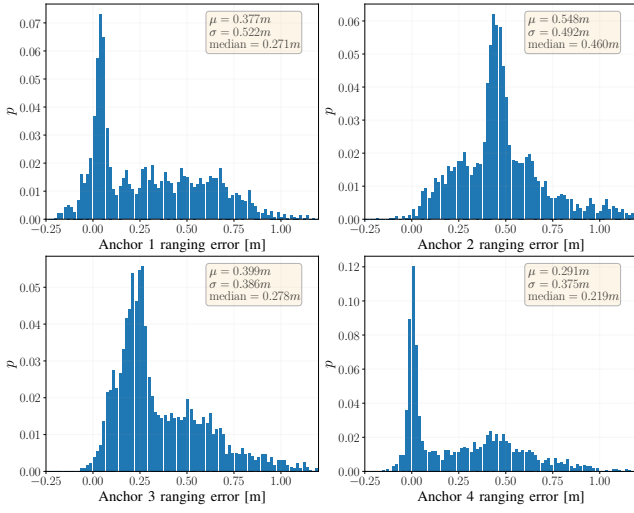

Fig. 3. Histogram of ranging errors for 4 anchors, also showing their mean μ, standard deviation σ and median.

which makes radio based ranging hard for estimation methods which assume gaussian noise. Errors are usually positive because radio waves can only get delayed by propagating through a medium with a propagation velocity lower than the speed of light or by traveling on a multipath when they are subject to reflections. The high mean ranging error of 0.55 m for anchor 2 can be only explained by module specific manufacturing tolerances or strong environment influences which lead to many NLOS measurements. Due to this asymmetry in the errors (especially for anchor 1, 3 and 4) a heavy-tailed distribution might be suitable.

Another problem are outliers whose values can be orders of magnitude higher than their real range value, on the right side of Fig. 2 it is visible that many outliers fall outside the 2 to 98 percentile range of the error distributions. Those outliers are likely caused by multipath effects and straight forward to filter out by employing a linear first or second order Kalman filter and thresholding the measurements against the current Mahalanobis distance [18]. For the sake of simplicity and the lack of additional scope we will still consider the measurement errors e_i of anchor i to be approximately gaussian distributed ($e_i \sim \mathcal{N}(\mu_i, \sigma^2)$) , in

which μ_i will be predicted by the error model for each tag-anchor pair and $\sigma = 0.35\,m$ is considered to be a fixed ranging measurement noise variance. This value is based on the ranging statistics found from several test runs, such as provided in Fig. 3.

C. Sensor error model

The objective is to design a model which predicts sensor ranging errors originating from the aforementioned error sources by taking into account the tag pose $\boldsymbol{p}_t(t)$, anchor poses $\boldsymbol{P}_a$ and ranging measurements $\boldsymbol{r}(t)$ as inputs. Directional antenna properties can be described by the antenna orientation and the LOS angles between tag and anchor positions using spherical coordinates, φ for the azimuth angle and θ for the inclination angle against the z-axis of the antenna frame (see Fig. 4). Because anchor antenna
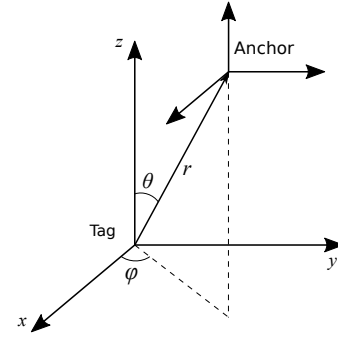

Fig. 4. Spherical coordinate parameterization of antenna LOS angles azimuth φ_t, inclination θ_t and range r between tag and anchor antenna coordinate frames, at the tag.

related errors are assumed to be constant and minimal for a static anchor placement without nearby objects, only the LOS angles φ_t and θ_t at the tag module are considered. In Fig. 4 one could imagine having the robot in Fig. 1 oriented towards an anchor in such a way that the LOS is blocked by robot components. An angle parameterization requires the desired error model to be periodic in those features. Normal polynomial basis functions such as used in [12] have shown poor performance because implicitly modelling periodicity in a given interval is not intuitive and non-trivial. Modelling arbitrary periodic functions can be accomplished using spline functions [19] with periodic boundary conditions. Another common model choice for modelling periodic functions are Fourier basis functions or trigonometric polynomials [20] which can be seen as a finite real-valued Fourier series (1), note that the bias term was removed in our notation.

$$f(x) = \sum_{m=1}^{M} a_m f_c(x, m) + b_m f_s(x, m) \qquad (1)$$

$$f_c(x, m) = \cos\left(\frac{2\pi m}{T}x\right) \qquad (2)$$

$$f_s(x, m) = \sin\left(\frac{2\pi m}{T}x\right) \qquad (3)$$

$$\text{with} \quad a_m, b_m, T \in \mathbb{R}$$

Here M denotes the amount of frequency components and T the period of the first harmonic. For our model

we treat all components of this series as single features and hence provide them as seperate components in (2) and (3). Additionally to the LOS angles, the measured distance $r(t)$ and a bias term are used in the feature matrix $\boldsymbol{X_f}$ in (5). To allow for higher order terms and interaction features, $\boldsymbol{X_f}$ is then mapped by $\boldsymbol{p}\colon \boldsymbol{X_f} \to \boldsymbol{X_p}$ into the polynomial feature matrix $\boldsymbol{X_p}$, consisting of all polynomial combinations of the features with degree less than or equal to P, e.g. mapping $[a, b]$ into degree-2 polynomial features would give $\left[a, b, a^2, ab, b^2\right]$. This function formulation leads to the following linear regression problem (4):

$$\boldsymbol{y} = \boldsymbol{X_p}\boldsymbol{\beta} + \boldsymbol{e}, \quad \hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \|\boldsymbol{y} - \boldsymbol{X_p}\boldsymbol{\beta}\| \tag{4}$$

$$\boldsymbol{X_f} = \Big[\mathbf{1}, \quad \boldsymbol{r}^T,$$
$$\boldsymbol{f_c}(\boldsymbol{\varphi_t}, 1;\ T = 2\pi)^T, \ldots, \boldsymbol{f_c}(\boldsymbol{\varphi_t}, M;\ T = 2\pi)^T,$$
$$\boldsymbol{f_s}(\boldsymbol{\varphi_t}, 1;\ T = 2\pi)^T, \ldots, \boldsymbol{f_s}(\boldsymbol{\varphi_t}, M;\ T = 2\pi)^T, \tag{5}$$
$$\boldsymbol{f_c}(\boldsymbol{\theta_t}, 1;\ T = \pi)^T, \ \ldots, \boldsymbol{f_c}(\boldsymbol{\theta_t}, M;\ T = \pi)^T,$$
$$\boldsymbol{f_s}(\boldsymbol{\theta_t}, 1;\ T = \pi)^T, \ \ldots, \boldsymbol{f_s}(\boldsymbol{\theta_t}, M;\ T = \pi)^T \Big]$$

Here $\boldsymbol{y}$ are the error values predicted by the sensor error model, $\boldsymbol{X_p}$ is the polynomial feature matrix and $\boldsymbol{\beta}$ the model parameters. After cross-validating models with different M and P values, the model with $M = 2$ and $P = 2$ gave the best bias-variance trade-off. This model parameterization produces 55 feature columns in matrix X_p. The fitted sensor
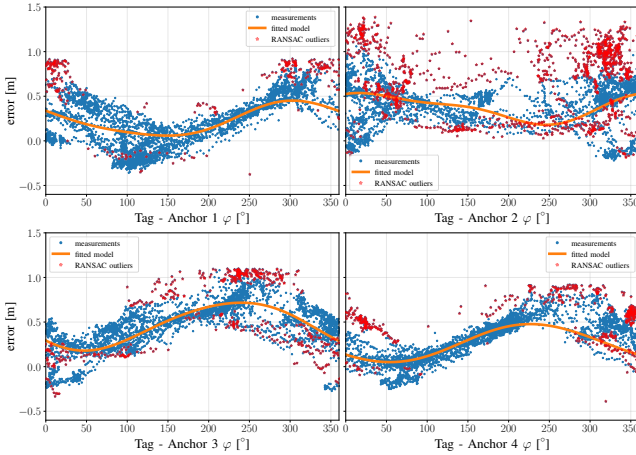

Fig. 5. Ranging errors and fitted sensor error model for 4 anchors in respect to only φ after removing outliers by comparing ranging measurments with estimated distances between tag and anchors by SLAM, according to (7).

error model for anchor 1 in Fig. 6 was evaluated between 0 and $360\,^\circ$ for φ and between 0 and $150\,^\circ$ for θ. The left side shows only the function response, the right side has the ground-truth training data overlayed for comparison, it's obvious that the covered training input range is narrow compared to possible values. However, values outside this range often are not applicable for most scenarios, where the θ range is small due to motion constraints in the roll and pitch axes which have the strongest influence on θ.
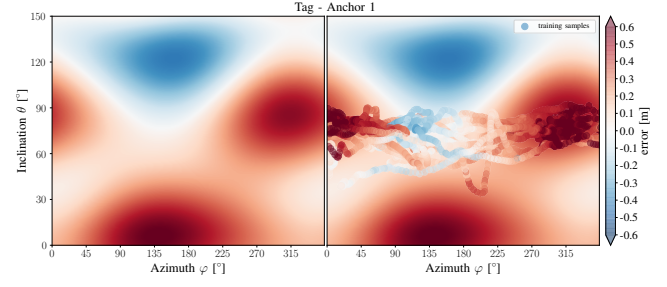

Fig. 6. Fitted sensor error model for anchor 1. Left: Visualized by evaluating it for $0 \le \varphi < 360\,^\circ$ and $0 \le \theta \le 150\,^\circ$. Right: Additionally overlayed with training samples. It is clearly visible that the training data only covers a small part of the model input range.

D. Calibration routine

During the calibration run only the visual navigation component is being used for localization, UWB ranging measurements are just logged away. Only after calibrating the sensor error model and estimating the anchor poses in the first step, ranging measurements are used in the navigation pipeline(Section III-C). The following steps outline our calibration routine:

1) Collect ranging measurements $\boldsymbol{r}(t)$ from tag module and pose estimates $\boldsymbol{P}(t) = \left[\boldsymbol{p}_{t_{xyz}}, \boldsymbol{q}_{t_{wxyz}}\right]^T$ while moving around in the area around the anchors making sure to cover many different tag poses for a well balanced dataset and map the anchor poses $\boldsymbol{P}_a = \left[\boldsymbol{p}_{a_{xyz}}, \boldsymbol{q}_{a_{wxyz}}\right]^T$ via AprilTag detections within the SLAM graph.
2) Compute pseudo-ranges $\tilde{\boldsymbol{r}}(t)$ by calculating the Euclidean norm between estimated tag and anchor position estimate: $\tilde{\boldsymbol{r}}(t) = \|\boldsymbol{p}_{t_{xyz}} - \boldsymbol{p}_{a_{xyz}}\|$.
3) Compute ranging error (6) and then select samples from $\boldsymbol{r}(t)$ into $\boldsymbol{r}'(t)$ by thresholding it by $\Delta\boldsymbol{r}(t) < 1\,m$ to filter out obvious outlier, see (7).
4) Calculate tag antenna LOS angles $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$ and assemble feature matrix $\boldsymbol{X_p}$.
5) Normalize feature matrix $\boldsymbol{X_p}$ and $\boldsymbol{r}'$ to have zero mean and unit variance to improve numerical stability and convergence of least-square solver in next step.
6) Fit model using a well regularized iterative linear least-square solver to prevent overfitting and RANSAC to drastically reduce the remaining outliers which passed the test in (7). We use ridge regression with an iterative conjugate gradient solver and a RANSAC scheme selecting at least $5\,\%$ samples for hypothesis testing to solve the linear regression problem in (4), where y equals $\boldsymbol{r}'(t)$.

In Fig. 5 we show the final fitted error model with respect to φ, while θ and r were kept constant. RANSAC recognized most of the remaining outliers very well, reducing the influence of outlier substantially.

$$\Delta\boldsymbol{r}(t) = |\boldsymbol{r}(t) - \tilde{\boldsymbol{r}}(t)| \tag{6}$$
$$\boldsymbol{r}'(t) = \{\boldsymbol{r}(t_i) \mid \Delta\boldsymbol{r}(t_i) < 1\,m\}, \forall t_i \tag{7}$$

III. Navigation pipeline

For local and global estimation, we use a loosely coupled SLAM system [21]. It is split into two parts: The first part is a local reference filter to fuse high-frequency data for fast robot state estimation. It is running at a frequency of 200 Hz to allow for attitude and position control under real-time constraints as required by highly dynamic systems such as MAVs. The second part is a global graph-based optimization that runs online but at a lower frequency, integrating filter estimates and landmark observations. This decoupled system allows to separate real-time critical estimation required for control and stabilization of an aerial system from slower but online global optimization needed to compensate drift of locally unobservable states. Our navigation pipeline receives

- gyrometer and accelerometer measurements at 500 Hz,
- delta 6D poses from a keyframe based visual odometry [22], [23], using 3D features at approx. 8 Hz,
- UWB ranging measurements from all anchors at approx. 30 Hz.

A. Local navigation filter

For high-frequency and jump-free state estimation suitable for control, we fuse IMU readings and visual odometry estimates in a local reference filter that is implemented as an error-state Extended Kalman Filter, described in detail in [24], [25]. It provides time-delay compensation for the stereo-based visual odometry, which provides relative transformations between keyframes [22]. The filter itself consists of two main parts: First a strapdown algorithm, which integrates angular rates and accelerations from the IMU and second the filter update step that processes the visual odometry estimates. Both, IMU measurements and visual odometry estimates, suffer from translational and rotational drift. The rotational drift can be partially compensated by stabilizing the roll and pitch angles. This is based on the assumption that, at least in the long term, the measured acceleration is dominated by gravity. Therefore, in a visual-inertial odometry system, the translation and rotation around the gravitational vector are globally not observable and the errors become unlimited over time [26]. This disadvantage can be compensated by adding sensors that perform an absolute measurement, such as UWB radio modules which measure distances between each other. Those distances are used in the non-linear multilateration equations (see Section III-C) for drift-free position estimation in x, y and z direction relative to a UWB coordinate system (u-frame). The reason this is not directly done in the Extended Kalman filter is the high recursive linearization error when dealing with ranging measurements [8]. The filter estimates the pose of a robot relative to a earth-fixed coordinate system (n-frame). In order to use the position measurements in the filter, the 6-dimensional transformation $^{n}T_{u}$ must be determined. We assume that at the beginning the UWB position values are accurate and we can directly set the translation of $^{n}T_{u}$ to the mean of the initially read UWB position values, while the system was static. Due to the chosen approach, the roll and tilt angles of the UWB coordinate system and the filter coordinate system match and can be set to 0 in the rotational part of $^{n}T_{u}$. The only undetermined variable of the transformation $^{n}T_{u}$ remains the rotation ϕ around the gravitational vector. It can be continuously estimated in the filter and extends the direct state x from [27] by the scalar quantity $^{n}\phi_{u}$

$$\boldsymbol{x} = \begin{bmatrix} ^{n}_{b}\boldsymbol{p}, ^{n}_{b}\boldsymbol{v}, ^{n}_{b}\boldsymbol{q}, ^{b}\boldsymbol{b}_{a}, ^{b}\boldsymbol{b}_{\omega}, ^{n}\phi_{u} \end{bmatrix}^{T} \quad (8)$$

where $^{n}_{b}\boldsymbol{p} \in \mathbb{R}^{3}$ is the position of the body frame (b-frame) relative to an earth-fixed, inertial frame (n-frame), $^{n}_{b}\boldsymbol{v} \in \mathbb{R}^{3}$ is the velocity, $^{n}_{b}\boldsymbol{q}$ the orientation represented as a quaternion and $^{b}\boldsymbol{b}_{a}$ and $^{b}\boldsymbol{b}_{\omega}$ are the acceleration and angular rate biases of the IMU. During the filter propagation the value of $^{n}\phi_{u}$ is not modified, but refined during the filter update. To minimize the influence of outliers in the UWB position readings a test based on the Mahalanobis distance [28] is performed and only measurements that passed are used in the filter.

B. Global pose and map estimation

We employ a graph SLAM system for global optimization of poses and 3D maps, which is based on the architecture presented in [29], [30], [21]. The online 3D mapping of the environment is based on the navigation filter estimates and dense depth data from Ardea's fisheye camera system described in [27]. Local maps of limited size and uncertainty, so-called submaps, are created by aggregating the merged depth data computed from four virtual pinhole stereo camera pairs along the trajectory estimated by the local reference filter. This is based on the assumption that the filter state estimates are locally stable but unobservable components such as position and the yaw angle can globally be subject to drift. The navigation filter being a local reference filter allows us to always switch its frame of reference into the origin of the current submap. This helps to maintain numerical stability and long-term consistency in the filter even in case of globally unbounded drift. It further allows for a better integration of the filter's estimates into the overlying SLAM graph according to its estimated uncertainties and probabilistic dependency structure without exposing filter-internal states to the overlying SLAM graph [29]. Submap origins are represented as nodes in the SLAM graph and connect via filter estimates as edges weighted by their respective Gaussian uncertainty. Artificial landmarks, in our case AprilTag markers [31] are attached to UWB anchors (see Fig. 8), constitute additional nodes. The detection and pose estimation of such landmarks result in 6D loop closure constraints, represented in the graph by edges between robot poses and landmarks, as visualized in Fig. 7. We then use incremental non-linear least-squares methods to compute global pose estimates for all nodes based on the sparse optimization problem defined by the SLAM graph.

C. UWB radio based localization

We employ the ToF based symmetric double-sided two-way-ranging (SDS-TWR) [15] scheme in order to not have
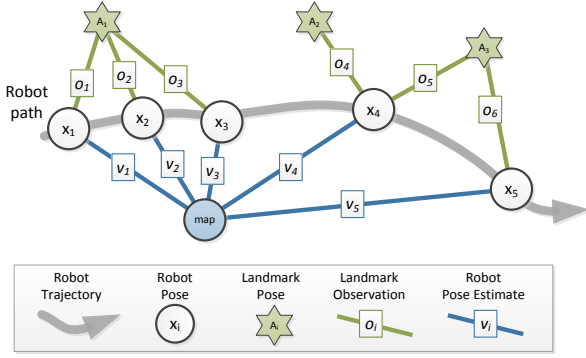
Fig. 7. Simplified visualization of SLAM graph, omitting the submapping component for clarity. Measurements are represented on edges, nodes represent hidden states which are estimated by SLAM.

to synchronize the clocks of all individual radio modules. Let $d_1, \cdots, d_N$ be the measured distances with associated errors $e_1, \cdots, e_N \sim \mathcal{N}(\mu, \sigma^2)$ estimated within the sensor error model in Section II. We are estimating the position of the tag $\hat{p}_t(t)$ at time t, given the previously measured static anchor positions $\boldsymbol{P}_A = [\boldsymbol{p}_{a1}, \cdots, \boldsymbol{p}_{a1}]^T$. Also those quantities are assumed to be Gaussian random variables. The multilateration equations are defined in (9) as

$$
\begin{aligned}
d_1(t_1) &= \|\boldsymbol{p}_{a1} - \hat{\boldsymbol{p}}_t(t_1)\| + e_1 \\
d_2(t_1) &= \|\boldsymbol{p}_{a2} - \hat{\boldsymbol{p}}_t(t_1)\| + e_2 \\
&\vdots \qquad \vdots \\
d_N(t_1) &= \|\boldsymbol{p}_{aN} - \hat{\boldsymbol{p}}_t(t_1)\| + e_N
\end{aligned}
\tag{9}
$$

Those equations are non-linear in the distance calculation between the tag $\boldsymbol{p}_t$ and the individual anchors $\boldsymbol{p}_{ai}$ using the Euclidean norm $\|\cdot\|_2$. As already mentioned in Section I-A and for the sake of brevity within the scope of this paper, using a non-linear least-squares solver is used as a universal and simple approach for solving such a non-linear estimation problem. In (10) the non-linear optimization objective function is given. Because squaring error terms containing outlier measurements strongly emphasize their influence on the optimization result, we use a loss function $\rho(z)$ which acts as a weighting function on the squared residuals to minimize the influence of outliers. We chose $\rho(z)$ to be the Cauchy loss function (11) which is considered to cope well with outliers [32].

$$
\hat{\boldsymbol{p}}_t = \arg\min_{\boldsymbol{p}_t} \sum_{i=1}^{N} \rho\left((d_i - \|\boldsymbol{p}_{Ai} - \boldsymbol{p}_t\|)^2\right)
\tag{10}
$$

$$
\rho(z) = \ln(z + 1)
\tag{11}
$$

The estimated tag position p_t and its uncertainty σ_{p_t} (propagated from constant $\sigma = 0.35\,m$) is then used in the update step of the local reference filter, described in Section III-A.

IV. EXPERIMENTAL SYSTEM SETUP

Our experimental MAV platform Ardea (see Fig. 1) is a hexacopter based on a triangular hexacopter frame construction propelled by 10" coaxial rotor pairs providing a

maximum thrust of 3.6 kg. It is designed to carry two pairs of fisheye stereo cameras without having any parts of the MAV in the field-of-view [27], [33]. The individual cameras have a field-of-view of approximately $80°$ horizontally and $125°$ vertically. Together the stereo camera setups have a $240°$ vertical field-of-view, which covers ground and ceiling and make it suitable for indoor environments where feature tracking on those surfaces provide a robust robot pose-estimate. The widely-used *Decawave DWM1000* UWB mod-

TABLE I
DECAWAVE DWM1000 UWB RANGING MODULE SETTINGS.

Channel number	2
Channel bandwidth	500 MHz
Pulse repetition frequency	64 MHz
Preamble length	1024 symbols
Data rate	6.8 Mbit/s
Output power	41.3 dBm/MHz

ules [11] are used for all ranging experiments and configured as provided in Table I. The tag module is mounted between the cameras at the front of the MAV as shown in Fig. 1. All anchor radio modules are equipped with AprilTag markers as shown in Fig. 8.
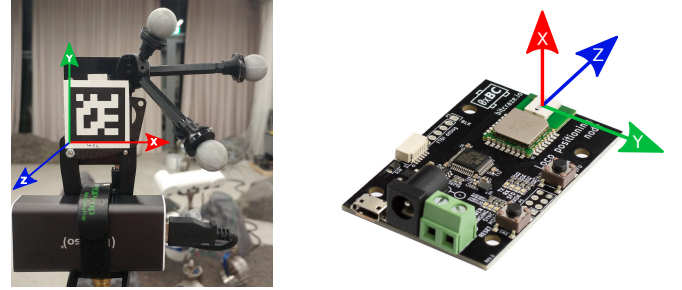

Fig. 8. UWB radio modules used by *Bitcraze* and coordinate frame conventions of AprilTags and UWB module antennas. Left: Backside with AprilTag. Right: Coordinate frame in UWB antenna frame.

V. EXPERIMENTS

In order to validate how well the anchor pose estimation and the sensor error model performs across different environment scenarios, experiments with two deviating anchor setups are conducted, see Fig. 10. Especially in regard to validating the sensor error model, those setups with different anchor positions and orientations were used as a means to provide different surroundings around the anchor antennas which alter the incoming and outgoing signals in a different way. The sensor ranging error, predicted by the sensor error model, and the anchor poses, estimated within the SLAM graph, are compared against ground-truth provided by a Vicon optical tracking system mounted on the ceiling of the lab. All anchor nodes as well as the multirotor platform are equipped with markers, see Fig. 8 and Fig. 1 respectively. Due to the complex experiment setup with the necessity to get a high variation of robot pose across the environment, the multirotor platform was carried around instead of putting effort into the generation of suitable trajectories.

The following three experiments are conducted:

1) Setup A with $N = 5$ anchors and a random path for calibration of anchor poses and the error model.
2) Anchor setup A as in 1 but a different path for evaluation how good the sensor error model performs at different tag positions. Therefore the error model parameters remain the same as in experiment 1.
3) Anchor setup B with $N = 5$ anchors, while only anchor poses are calibrated and updated again, the sensor error model still remains the same as in experiment 1. This experiment should show how well the error model generalizes across different anchor setups.
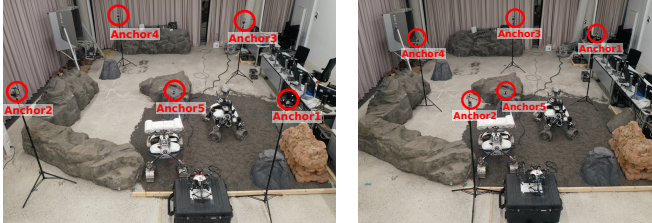

Fig. 9. Left: Anchor setup A. Right: Anchor setup B. Note that not only the positions, but also orientations changed between both setups.

After processing the logged dataset, the anchor poses estimated by the SLAM system can be extracted instantly. In Fig. 10 the anchor mapping result for experiment 1 is shown and Table II lists the associated anchor position errors for all 3 experiments, in which anchor positions were re-calculated every time. Those errors are mainly caused by the inherent drift of the VINS, however due to regular occurring SLAM loop closures upon re-visiting AprilTag landmarks, this drift is reduced substantially. The AprilTag pose estimation accuracy itself as well as the complexity of the randomly chosen robot trajectories have an impact on this error too. Faster robot movements might have lead to noticeable motion blur and therefore degraded accuracy in the feature tracking of the VINS. The trajectory in experiment 1 resulted in the highest anchor pose estimation error of 11.4 cm root-mean-square error (RMSE), which is however still below anchor localization errors of 59.2 cm and 15.2 cm mentioned in [34] and [35] respectively, which are solely based on range measurements.

Table III shows how the sensor error model performed across the experiments. For brevity only results for 4 anchors are shown. As result metrics we used the R2 score and the RMSE. The R2 score, or also known as coefficient of determination, which ranges from 0 to 1, 0 being the worst and 1 indicating that the regression predictions perfectly fit the data. It is a suitable indicator of how well variance is explained by the model and well it generalizes to new inputs. The RMSE is provided for the experiments without applying the error model calibration and with the calibration applied by subtracting the predicted error by the real error measured by the optical tracking system ground-truth. It is not surprising that the first experiment, which was used for fitting the error model, showed the best results. Interestingly the R2 score of anchor 1 in experiment 1 is slightly lower
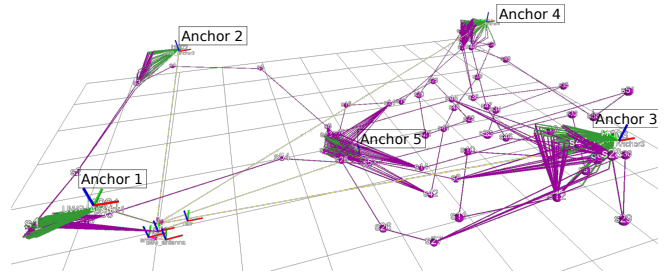

Fig. 10. Estimated anchor poses of setup A in SLAM graph visualization. Coordinate frames represent ground-truth measurements by the optical tracking system, the error are listed in Table II. The short green lines visualize the 6D pose constraint edges between tag and landmark positions at tag positions where AprilTag detections were made. The remaining purple lines are SLAM graph edges connecting those tag positions with the current submap origin and also between submaps, all the way to the map world frame.

TABLE II

RESULTING ERRORS OF ESTIMATED ANCHOR POSITIONS BY THE SLAM SYSTEM FOR 4 ANCHORS AND ALL 3 EXPERIMENTS.

		1	2	3
Anchor 1	x [mm]	24	54	-58
	y [mm]	-68	15	-140
	z [mm]	-33	-47	-66
Anchor 2	x [mm]	-116	-4	-17
	y [mm]	-106	146	-21
	z [mm]	-50	-9	-14
Anchor 3	x [mm]	206	65	32
	y [mm]	133	91	-132
	z [mm]	3	-32	-68
Anchor 4	x [mm]	120	56	92
	y [mm]	210	77	-125
	z [mm]	62	-39	2
Total RMSE	[mm]	114	70	79

than in the other experiments, which might be an indicator of robot and anchor poses which are not well captured by the sensor error model. The sensor error model also shows good performance in experiment 2. In experiment 3, where the anchors have been placed in different locations the error model still makes reasonable predictions which proves that the main influence on the total error seems to originate from the immediate environment of the tag antenna and less from the anchor antennas.

VI. CONCLUSION AND FUTURE WORK

In our experiments we have shown an elegant workflow how to jointly calibrate anchor poses as well as an efficient linear sensor error model in one calibration data collection run. The SLAM system estimates the anchor poses already during the data collection online, in a second offline step the sensor error model parameters are estimated. Because visual navigation is very accurate over short distances it proved to be suitable for anchor position and sensor error model calibration of UWB ranging modules which improved the ranging error substantially. Moreover, using AprilTags as artificial landmarks on the UWB anchors helps getting many loop-closures, which reduces the drift in the VINS. In future work we want to investigate and compare more elaborate

		1*	2	3
Anchor 1	R2 score	0.74	0.81	0.81
	RMSE w/o [mm]	132	109	151
	RMSE with [mm]	11	14	16
Anchor 2	R2 score	0.63	0.34	0.23
	RMSE w/o [mm]	307	315	332
	RMSE with [mm]	33	54	55
Anchor 3	R2 score	0.79	0.67	0.74
	RMSE w/o [mm]	143	240	139
	RMSE with [mm]	14	24	17
Anchor 4	R2 score	0.79	0.73	0.42
	RMSE w/o [mm]	143	127	204
	RMSE with [mm]	14	17	40

* dataset used for fitting the sensor error model

sensor error models which benefit from the estimated anchor orientations and also predict a meaningful error variance for a complete uncertainty propagation into the navigation framework. It would also be beneficial to omit this additional calibration step and instead have an online joint estimation scheme for both, sensor error model as well as anchor poses, which then only activates ranging based localization once anchor poses as well as sensor error models are considered sufficiently accurate by a calibration quality metric.

VII. ACKNOWLEDGMENT

REFERENCES

[1] T. Tomic *et al.*, "Toward a fully autonomous UAV," *IEEE Robotics & Automation Magazine*, vol. 19, no. 3, pp. 46–56, 2012.

[2] K. Schmid *et al.*, "Autonomous vision-based micro air vehicle for indoor and outdoor navigation," *Journal of Field Robotics*, vol. 31, no. 4, pp. 537–570, Jul. 2014.

[3] J. Delmerico and D. Scaramuzza, "A benchmark comparison of monocular visual-inertial odometry algorithms for flying robots," in *ICRA*, May 2018.

[4] J. J. S. Jr., P. Axelrad, B. W. Parkinson, and P. Enge, "Satellite constellation and geometric dilution of precision," *Global Positioning System: Theory and Applications, Progress in Astronautics and Aeronautics*, vol. I, pp. 177–208, 1994.

[5] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 37, no. 6, pp. 1067–1080, Nov 2007.

[6] N. Sirola, "Closed-form algorithms in mobile positioning: Myths and misconceptions," in *2010 7th Workshop on Positioning, Navigation and Communication*, March 2010, pp. 38–44.

[7] J. A. J. Smith, "Closed-form least-squares source location estimation from range-difference measurements," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1987.

[8] J. Choliz, A. Hernandez-Solana, and A. Valdovinos, "Evaluation of algorithms for uwb indoor tracking," in *Positioning Navigation and Communication (WPNC), 2011 8th Workshop on*, April 2011.

[9] L. Zwirello *et al.*, "Uwb localization system for indoor applications: Concept, realization and analysis," *JECE*, 2012.

[10] A. Ledergerber and R. D'Andrea, "Ultra-wideband range measurement model with gaussian processes," in *CCTA*, Aug 2017, pp. 1929–1934.

[11] Decawave, *DWM1000 Datasheet version 1.7*, Decawave, 2016, Accessed 12 Jan. 2019.

[12] M. Stampa, M. Müller, D. Heß, and C. Röhrig, "Semi-automatic calibration of uwb range measurements for an autonomous mobile robot," in *ISR*, June 2018.

[13] H. Wymeersch, S. Marano, W. M. Gifford, and M. Z. Win, "A machine learning approach to ranging error mitigation for uwb localization," *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1719–1728, June 2012.

[14] A. Ledergerber and R. D'andrea, "Calibrating away inaccuracies in ultra wideband range measurements: A maximum likelihood approach," *IEEE Access*, vol. 6, pp. 78 719–78 730, 2018.

[15] R. Hach, *Symmetric double side two way ranging*, Contribution to IEEE 802.15.4A, IEEE P802.15 Working Group for Wireless Personal Area Networks (WPAN), IEEE Computer Society NY, USA Std., 2005.

[16] Decawave, *Antenna delay calibration of DW1000-based products and systems, version 1.2, Appl. Note APS014*, Decawave, 2018, Accessed 12 Jan. 2019.

[17] J. Volakis, *Antenna Engineering Handbook, Fourth Edition*. McGraw-Hill Professional, 2007.

[18] H. Wu, S. Chen, B. Yang, and K. Chen, "Feedback robust cubature kalman filter for target tracking using an angle sensor," *Sensors*, vol. 16, no. 5, p. 629, 2016.

[19] C. De Boor, *A practical guide to splines; rev. ed.*, ser. Applied mathematical sciences. Berlin: Springer, 2001.

[20] M. Powell, *Approximation Theory and Methods*. Cambridge University Press, 1981.

[21] M. J. Schuster, K. Schmid, C. Brand, and M. Beetz, "Distributed stereo vision-based 6d localization and mapping for multi-robot teams," *Journal of Field Robotics*, Oct. 2018.

[22] H. Hirschmüller, P. R. Innocent, and J. M. Garibaldi, "Fast, unconstrained camera motion estimation from stereo without tracking and robust statistics," in *ICARCV*, 2002.

[23] A. Stelzer, H. Hirschmüller, and M. Görner, "Stereo-vision-based navigation of a six-legged walking robot in unknown rough terrain," *IJRR*, vol. 31, no. 4, pp. 381–402, 2012.

[24] K. Schmid *et al.*, "State estimation for highly dynamic flying systems using key frame sodometry with varying time delays," in *IROS*, 2012.

[25] K. Schmid, F. Ruess, and D. Burschka, "Local reference filter for life-long vision aided inertial navigation," in *17th FUSION*, 2014.

[26] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, "Camera-imu-based localization: Observability analysis and consistency improvement," *The International Journal of Robotics Research*, vol. 33, no. 1, pp. 182–201, 2014.

[27] M. G. Müller *et al.*, "Robust visual-inertial state estimation with multiple odometries and efficient mapping on an mav with ultra-wide fov stereo vision," in *IROS*, 2018.

[28] F. M. Mirzaei and S. I. Roumeliotis, "A kalman filter-based algorithm for imu-camera calibration: Observability analysis and performance evaluation," *IEEE transactions on robotics*, 2008.

[29] M. J. Schuster *et al.*, "Multi-robot 6D graph SLAM connecting decoupled local reference filters," in *IROS*, 2015.

[30] M. J. Schuster *et al.*, "Towards autonomous planetary exploration," *Journal of Intelligent & Robotic Systems*, 2017.

[31] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *IEEE International Conference on Robotics and Automation*, 2011.

[32] J. T. Barron, "A more general robust loss function," *CoRR*, vol. abs/1701.03077, 2017. [Online]. Available: http://arxiv.org/abs/1701.03077

[33] P. Lutz *et al.*, "ARDEA - An MAV with skills for future planetary missions," *Journal of Field Robotics*, 2019, manuscript submitted for publication.

[34] K. Batstone, M. Oskarsson, and K. Astrom, "Robust time-of-arrival self calibration and indoor localization using wi-fi round-trip time measurements," in *2016 IEEE International Conference on Communications Workshops (ICC)*, May 2016, pp. 26–31.

[35] K. Batstone, M. Oskarsson, and K. Astrom, "Towards real-time time-of-arrival self-calibration using ultra-wideband anchors," in *Proccedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2017.