

Scalable and Modular Ultra-Wideband Aided Inertial Navigation

Roland Jung¹ and Stephan Weiss¹

Abstract— Navigating accurately in potentially GPS-denied environments is a prerequisite of autonomous systems. Relative localization based on ultra-wideband (UWB) is – especially indoors – a promising technology.

In this paper, we present a probabilistic filter based Modular Multi-Sensor Fusion (MMSF) approach with the capability of using efficiently all information in a fully meshed UWB ranging network. This allows an accurate mobile agent state estimation and the calibration of the ranging network’s spatial constellation. We advocate a new paradigm that includes elements from Collaborative State Estimation (CSE) and allows us considering all stationary UWB anchors and the mobile agent as a decentralized set of estimators/filters. With this, our method can include all meshed (inter-)sensor observations tightly coupled in a modular estimator. We show that the application of our CSE-inspired method in such a context breaks the computational barrier. Otherwise, it would, for the sake of complexity-reduction, prohibit the use of all available information or would lead to significant estimator inconsistencies due to coarse approximations. We compare the proposed approach against different MMSF strategies in terms of execution time, accuracy, and filter credibility on both synthetic data and on a dataset from real Unmanned Aerial Vehicles (UAVs).

I. INTRODUCTION AND RELATED WORK

Accurate localization is a crucial component of autonomous robotic systems, e.g., service robots, warehouse pallet robots, etc., to control and navigate accurately in unknown and Global Positioning System (GPS)-denied environments. Especially indoors, the demand for a cost friendly, scalable and accurate positioning infrastructure is growing. Typically, range-based localization systems use time of arrival (TOA), time difference of arrival (TDOA), return time of flight (RTOF) [1] or received signal strength (RSS) metrics/lateration techniques to estimate the distance between the antennas [2]. A promising technology for both data transmission and localization is based on ultra-wideband (UWB) radio frequency (RF) signals [2]. It has desirable features for estimating distances between two transceivers, as its large bandwidths allow the UWB receiver to accurately estimate the arrival time of the first signal path [2].

For obtaining a unique 3-degrees of freedom (DoF) position estimate of a moving UWB device/node (tag), a classical multilateration approach requires simultaneous range measurements of at least four known stationary UWB nodes (anchors). The precision of the multilateration is strongly

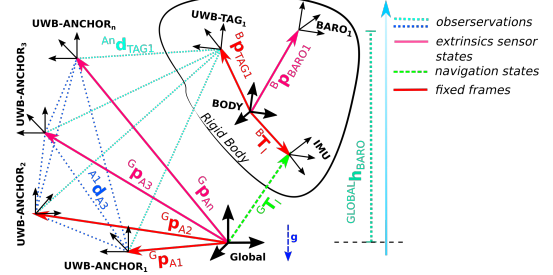


Fig. 1: Spatial frame constellation of the proposed UWB inertial aided navigation framework.

dependent on (i) the relative pose between the ranging modules’ antennas [3], (ii) the placement of stationary modules (anchors) as it influences the positional dilution of precision (PDOP) of the mobile modules, and (iii) multipath effects and non-line-of-sight (NLOS) condition may lead to wrong distance estimates [4]. Therefore, it is reasonable to combine these range measurements of mobile modules with other complementary sensors, e.g., an Inertial Measurement Unit (IMU) and a barometer, as depicted in Figure 1, to increase the robustness against dropouts, precision, and accuracy [5]–[10].

Hol et al. show in [4], that tightly fusing range measurements with inertial measurements obtained by an IMU in a probabilistic fusion algorithm allows (i) to recover a 6-DoF pose, (ii) to bridge periods with limited UWB range measurements, and (iii) to successfully detect and reject outliers. They assumed known and static anchor locations, which may cause unmodeled errors if the anchor positions were not accurately measured initially. However, calibration routines to estimate the position of deployed anchors have been presented in, e.g., [5], [6], [8], [11]–[14].

In [12], Hausman et al. proposed a multi-sensor fusion approach combining inertial sensor data with loosely coupled vision-based pose measurements, GPS measurements, and tightly coupled UWB range measurements for precision landing. They proposed a Linear Least Squares (LLS) initialization scheme for anchors, based on estimated tag positions, that provides the estimator with an initial belief for three estimated anchor positions.

In recent years, relative state estimation approaches to eliminate the need for stationary anchors in GPS-denied environment have been presented in [15]–[19]. For instance, Guo et al. present in [15], [17] an infrastructure-free co-operative approach to estimate the positions of neighboring Unmanned Aerial Vehicles (UAVs). Similarly, Xu et al. presented in [16] an optimization-based, fully decentralized visual-inertial-UWB fusion framework for relative state estimation in a swarm of UAVs. Nguyen et al. extended

¹Roland Jung, and Stephan Weiss are with the Department of Smart Systems Technologies in the Control of Networked Systems Group, University of Klagenfurt, Austria {roland.jung, stephan.weiss}@iieee.org

This work was partially supported by the EU-H2020 project BUG-WRIGHT2 (GA 871260) and from the Austrian Ministry of Climate Action and Energy (BMK) under the grant agreement 891124 (RoMInG).

Pre-print version, accepted June/2022, DOI follows ASAP ©IEEE.

in [9] a state-of-the-art optimization-based Visual-Inertial Odometry (VIO) algorithm to fuse camera, IMU and UWB range measurements from a single anchor to reduce the drift. Having only a single anchor in this configuration allows rendering the relative position with respect to the navigation frame observable, but the orientation about the gravity vector is still unobservable. Therefore, at least two anchors need to be known to achieve a fully observable system. Song et al. fused in [7], Light Detection and Ranging (LiDAR), UWB, and inertial measurements in an EKF-Simultaneous Localization and Mapping (SLAM) algorithm, while the UWB ranging measurement reduced accumulated errors in the proposed LiDAR-based SLAM algorithm.

Shi et al. investigated recently in [8] on the anchor self-calibration in a tightly-coupled UWB-ranging and IMU fusion algorithm in simulations with five anchors and varying ranging noise. In the first step of this algorithm, a coarse anchor position initialization, minimizing the residual between predicted tag positions and unknown anchor position in a least-squares problem. In the second step, these initial guesses are used for estimating the full state in a regular Error-State Extended Kalman Filter (ESEKF) [20] heavily limiting the scalability of the approach to more anchors. Also, the previously mentioned approaches suffer either from poor scalability or coarse approximations, affecting estimator consistency [21].

Therefore, we revisit the infrastructure-based UWB inertial localization and continuous/online anchor self-calibration for large UWB networks using on our recently proposed Modular Multi-Sensor Fusion Decoupled Approximated History (MMSF-DAH) approach [22]. It is based on EKF decoupling strategies originating from Collaborative State Estimation (CSE) [23] and has promising attributes that can be extended to render real-time tightly coupled and scalable UWB aided inertial navigation. From that work, we borrow the idea to treat any inter-sensor observation in an isolated fashion, requiring only the participating sensor estimates. For our present approach, by associating each UWB device to a single sensor estimator in the MMSF framework, measurements between any UWB device can be processed efficiently. This step allows combining the modular inclusion of states of external sensors with a regular MMSF with onboard sensors. Our main contributions can be summarized as:

- We propose a modular and scalable UWB-inertial based ESEKF, merging aspects from CSE and MMSF in order to estimate an agent's 6-DoF motion and sensor calibration states, and simultaneously estimate in a SLAM-like fashion geometry-states of very large UWB sensor networks in real-time.
- We perform extensive Monte Carlo simulations on synthetic and on real data from a UAVs dataset to verify both consistency and accuracy of the proposed approach.
- We evaluate the self-calibration of simulated anchor positions using different sensor configurations and against different state-of-the-art fusion strategies, showing im-

proved performance by incorporating meshed range measurements between stationary UWB anchors.

A. Notation

Throughout this paper, the following notation is used. A normally distributed multivariate variable is defined as $\mathbf{X}_i \sim \mathcal{N}(\hat{\mathbf{x}}_i, \Sigma_{ii})$, with a mean $\hat{\mathbf{x}}_i$ and covariance (uncertainty) Σ_{ii} , which is called the *belief* of i . The time indices of state variables are indicated by the right superscript. E.g. \mathbf{X}^k , denotes the state at the time t^k . Names of reference frames are capitalized and calligraphic, e.g. \mathcal{I} for IMU. A pose between the reference frames \mathcal{A} and \mathcal{B} is defined as ${}^{\mathcal{A}}\mathbf{T}_{\mathcal{B}} \in SE(3) := \left\{ \begin{bmatrix} {}^{\mathcal{A}}\mathbf{R}_{\mathcal{B}} & {}^{\mathcal{A}}\mathbf{p}_{\mathcal{B}} \\ \mathbf{0}^{\top} & 1 \end{bmatrix} \mid \mathbf{R} \in SO(3), \mathbf{p} \in \mathbb{R}^3 \right\}$ (read as ${}^{from}_{to} \mathbf{x}$). The operators \oplus and \ominus should emphasize that state vector elements need to be treated differently. The error of translational estimates are additive, e.g. ${}^{\mathcal{G}}\mathbf{p}_{\mathcal{I}} = {}^{\mathcal{G}}\hat{\mathbf{p}}_{\mathcal{I}} + {}^{\mathcal{G}}\tilde{\mathbf{p}}_{\mathcal{I}}$, while rotational errors are right-multiplicative, e.g. ${}^{\mathcal{G}}\mathbf{R}_{\mathcal{I}} = {}^{\mathcal{G}}\hat{\mathbf{R}}_{\mathcal{I}}(\mathbf{I}_3 + [{}^{\mathcal{G}}\tilde{\boldsymbol{\theta}}_{\mathcal{I}}]_{\times}) \in SO^3$. (\bullet) specifies an estimate, $(\tilde{\bullet})$ an error, and $(\bar{\bullet})$ an average. The z-axis of the global navigation reference frame is gravity aligned. $(\mathbf{T})_{\{x,y,z,\mathbf{p},\mathbf{R}\}}$ and $(\mathbf{p})_{\{x,y,z\}}$ refers to the corresponding element.

II. MODULAR ULTRA-WIDEBAND INERTIAL AIDED NAVIGATION

In the proposed sensor constellation, an IMU is used as a proprioceptive state propagation sensor, a barometer is used for a tightly coupled height estimation (both sensors are typically available on Micro Aerial Vehicle (MAV)) and a single UWB tag that performs ranging measurements with the UWB anchors in communication range. Also, UWB anchors perform measurements among them if in range.

We assume commercial UWB modules to work with the double-sided two-way-ranging protocol (High Precision Ranging (HiPR)) proposed by Neuhold et al. [1] allowing fast ranging acquisitions at a rate of 40 Hz by performing a Round-Robin scheduling to avoid network congestion, and measurement broadcasting to close-by sensor nodes. The UWB tag acquires/sniffs range measurement to and between nearby stationary UWB anchor upon fly-by in communication range and fuses them locally in a probabilistic modular filter framework, while, in contrast to other work, consistently accounting for correlations between individual sensor estimates. For simplicity, other than Gaussian noise, we do not assume any biases on the signals, nor extrinsic calibration between tag and IMU. We refer to [14] on how to include these elements into an UWB-inertial estimator and how to perform a priori a coarse anchor position initialization.

Each sensor has a unique identifier and belongs to a known class of sensor types. Measurements are processed in a ESEKF-based MMSF framework [22], which associates each measurement to a sensor estimator instance that performs the information fusion with potentially other sensors of the sensor suite. Each sensor estimator handles sensor specific states, such as calibration parameters needed for self-calibration. Further, each sensor estimator performs a

statistical Normalized Innovation Squared (NIS) [21] hypothesis check to detect and reject outliers, and is capable of processing delayed (out-of-order) measurements (see [22]).

Applying joint observations, incorporating different estimates, results in cross-covariance terms between them and at some point, all estimators might be correlated. Our formulation allows treating different subsets of these state variables from a CSE perspective of separate estimators (i.e., entities).

Classical centralized filters such as [12] are typically not truly modular, as they operate on the predefined full state as one entity (including the full joint covariance matrix) making individual sensor filter step very costly in large networks.

We integrate our approach in three different MMSF strategies (cf. [22] for details on the general concepts of these approaches) and compare the resulting performance: MMSF-C is a centralized-equivalent EKF filter implementation performing all filter steps on the entire full state vector. MMSF-DP is a centralized-equivalent EKF filter implementation, performing the state propagation of individual sensor-estimators independently, while update steps are performed on the full state vector. MMSF-DAH performs all filter steps isolated, requiring only sensor-estimators that are directly involved (so-called participants) in the filter update steps, while correlations to non-participants are conservatively approximated. We show in Section III that this last MMSF strategy is best suited for our real-time UWB-inertial aided navigation and mesh self-calibration approach.

A. Sensor Suite

The sensor suite consists of a varying number of stationary UWB anchors, an UWB tag, a barometer, and an IMU. In the following, we describe the corresponding measurement models and state definitions.

Due to indirect error estimation [20], observations have to be expressed by their error $\tilde{\mathbf{z}} = \mathbf{z} \ominus \hat{\mathbf{z}}$. This measurement error needs to be linearized with respect to the error state at the current estimate $\tilde{\mathbf{z}} = \mathbf{H}\tilde{\mathbf{x}}$ with the measurement Jacobian $\mathbf{H} = \left. \frac{\partial h}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}}$ for the measurement function $\mathbf{z} = h(\mathbf{x})$. Due to space constraints, we omit the details of the measurement Jacobians and the error-state definition.

1) *Inertial Measurement Unit:* We use the IMU as state propagation sensor in an indirect filter formulation [24]

$$\mathbf{X}_{\mathcal{I}} = [\mathcal{G}\mathbf{p}_{\mathcal{I}}, \mathcal{G}\mathbf{v}_{\mathcal{I}}, \mathcal{G}\mathbf{q}_{\mathcal{I}}, \mathcal{I}\mathbf{b}_{\omega}, \mathcal{I}\mathbf{b}_a]^T \quad (1)$$

with $\mathcal{G}\mathbf{p}_{\mathcal{I}}, \mathcal{G}\mathbf{v}_{\mathcal{I}}$, and $\mathcal{G}\mathbf{q}_{\mathcal{I}}$ as the position, velocity, and orientation of the IMU \mathcal{I} w.r.t. to the global navigation frame \mathcal{G} . $\mathcal{I}\mathbf{b}_{\omega}$ and $\mathcal{I}\mathbf{b}_a$ are the estimated gyroscope and accelerometer biases to correct the related IMU readings. $\mathcal{B}\mathbf{p}_{\mathcal{I}}$ and $\mathcal{B}\mathbf{q}_{\mathcal{I}}$ are the constant (and thus not included in the state vector) values for position and orientation of the IMU \mathcal{I} referring to the body frame \mathcal{B} . The nominal-state kinematics

are described by

$$\begin{aligned} \mathcal{G}\dot{\mathbf{p}}_{\mathcal{I}} &= \mathcal{G}\mathbf{v}_{\mathcal{I}}, \\ \mathcal{G}\dot{\mathbf{v}}_{\mathcal{I}} &= \mathcal{G}\mathbf{R}_{\mathcal{I}}(\mathcal{I}\mathbf{a}^{\#} - \mathbf{b}_a - \mathbf{n}_a) + \mathcal{G}\mathbf{g}, \\ \mathcal{G}\dot{\mathbf{R}}_{\mathcal{I}} &= \mathcal{G}\mathbf{R}_{\mathcal{I}}[\mathcal{I}\boldsymbol{\omega}^{\#} - \mathbf{b}_{\omega} - \mathbf{n}_{\omega}]_{\times}, \\ \dot{\mathbf{b}}_a &= \mathbf{n}_{b_a}, \dot{\mathbf{b}}_{\omega} = \mathbf{n}_{b_{\omega}}, \end{aligned} \quad (2)$$

where $\mathcal{I}\mathbf{a}^{\#}$ and $\mathcal{I}\boldsymbol{\omega}^{\#}$ are the accelerometer and gyroscope measurements of the IMU with a white measurement noise \mathbf{n}_a and \mathbf{n}_{ω} . $\mathbf{n}_{b_a} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{b_a})$ and $\mathbf{n}_{b_{\omega}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{b_{\omega}})$ are assumed to be white Gaussian noise to model the bias change over time as a random process. The gravity vector is assumed to be aligned with the z-axis of the navigation frame $\mathcal{G}\mathbf{g} = [0, 0, 9.81]^T$.

The IMU provides noisy and biased linear acceleration and angular velocity samples in the sensor frame \mathcal{I} .

$$\mathbf{z}_a = \mathcal{I}\mathbf{a}^{\#} = \mathcal{I}\mathbf{a} + (\mathcal{G}\mathbf{R}_{\mathcal{I}})^T \mathcal{G}\mathbf{g} + \mathcal{I}\mathbf{b}_a + \mathcal{I}\mathbf{n}_a, \quad (3)$$

$$\mathbf{z}_{\omega} = \mathcal{I}\boldsymbol{\omega}^{\#} = \mathcal{I}\boldsymbol{\omega} + \mathcal{I}\mathbf{b}_{\omega} + \mathcal{I}\mathbf{n}_{\omega}, \quad (4)$$

with a Gaussian accelerometer $\mathcal{I}\mathbf{n}_a \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_a)$ and gyroscope $\mathcal{I}\mathbf{n}_{\omega} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\omega})$ noise.

2) *Pressure sensor:* We include a barometric pressure sensor to show-case the versatility of our approach, not being limited to IMU and UWB measurement inclusion but also able to treat additional sensors as separate estimators in the CSE spirit. The barometer measures the local absolute pressure. Its position $\mathcal{B}\mathbf{p}_{\mathcal{P}}$ with respect to the body reference frame is assumed to be fixed and known, and is not included in the estimation process.

For the relative height estimation based on the barometric formula of the isotherm atmosphere model [25], a reference pressure $\mathcal{G}P$ and temperature $\mathcal{G}T$ at the global reference frame \mathcal{G} needs to be known (e.g., by a sensor on the ground). The barometric pressure $\mathcal{P}P$ measured by the on-board sensor is then:

$$\mathcal{P}P = \mathcal{G}P \left(1 - \frac{c_1 \mathcal{G}h_{\mathcal{P}}}{\mathcal{G}T} \right)^{c_2}. \quad (5)$$

The standard atmosphere at sea-level on the equator specifies a constant temperature lapse rate of $c_1 = 0.0065$ K/m valid up to 11 km and the constant $c_2 = 5.257$ for dry air. The estimated height of the barometer expressed by the states is

$$\mathcal{G}h_{\mathcal{P}} = (\mathcal{G}\mathbf{T}_{\mathcal{P}})_z = \left(\mathcal{G}\mathbf{T}_{\mathcal{I}} \mathcal{B}\mathbf{T}_{\mathcal{I}}^{-1} \mathcal{B}\mathbf{T}_{\mathcal{P}} \right)_z. \quad (6)$$

Inserting Equation (6) in Equation (5) yields the noisy pressure reading modeled as

$$\mathcal{G}z_{\mathcal{P}} = \mathcal{P}P + n_{\mathcal{P}}, \quad (7)$$

with Gaussian noise $n_{\mathcal{P}} \sim \mathcal{N}(0, \sigma_{\mathcal{P}}^2)$.

3) *Range sensor:* To model the range measurements between two UWB nodes, e.g., between tag and anchor or between anchors, we include the 3D position of each anchor $\mathcal{G}\mathbf{p}_A$ in the state estimation process.

$$\mathbf{X}_{\mathcal{A}} = [\mathcal{G}\mathbf{p}_A], \quad (8)$$

which does not change over time:

$$\dot{\mathbf{g}}_{\mathcal{B}} \mathbf{p}_{\mathcal{A}} = \mathbf{0}. \quad (9)$$

The transformation between the body reference frame \mathcal{B} and the UWB tag $\mathcal{B}_{\mathcal{T}}$ is assumed to be fixed and not included in the estimation. The range measurement between two anchors (A2A), or between a tag and an anchor (T2A) is modelled as

$$\{\mathcal{A}, \mathcal{T}\} z_{\{\mathcal{A}, \mathcal{A}\}} = \{\mathcal{A}, \mathcal{T}\} d_{\{\mathcal{A}, \mathcal{A}\}} + n_{\text{UWB}}, \quad (10)$$

with Gaussian noise $n_{\text{UWB}} \sim \mathcal{N}(0, \sigma_{\text{UWB}}^2)$ and

$$\tau d_{\mathcal{A}} = \left\| \mathbf{g}_{\mathcal{B}} \mathbf{p}_{\mathcal{A}} - \mathbf{g}_{\mathcal{B}} \mathbf{p}_{\mathcal{T}} \right\|_2, \quad (11)$$

$$\mathcal{A}_i d_{\mathcal{A}_j} = \left\| \mathbf{g}_{\mathcal{B}} \mathbf{p}_{\mathcal{A}_j} - \mathbf{g}_{\mathcal{B}} \mathbf{p}_{\mathcal{A}_i} \right\|_2. \quad (12)$$

4) *Estimator Summary*: Summarized and as depict in Figure 1, our modular ESEKF formulation is based on the IMU navigation states and their dynamics in Equation (2). The filter obtains corrections from the barometer (Equation (5) and Equation (7)), the measurements between UWB tag and anchors in range (Equation (11)) as well as the inter-anchor measurements observed from anchors in range (Equation (12)). The inclusion of the inter-anchor measurements as an extension to our CSE inspired MMSF framework ([22]) is key to enable consistent and scalable inclusion of the UWB mesh geometry in the estimation process in real-time. This renders our approach a real-time capable consistent UWB-inertial SLAM-like estimator.

III. EVALUATIONS

The experiments are done in our MATLAB framework, that allows to load real data from the EuRoC [26] dataset or to generate smooth trajectories and noisy, biased IMU samples. Exteroceptive measurements are generated based on the ground truth trajectory and are modified by the sensors' calibration states and noise parameters. Furthermore, a delay and dropout rate can be applied to these measurements. The noisy and biased real-world IMU samples provided by the datasets are used without modifications. Finally, in multiple Monte Carlo simulation runs, all measurements are processed in a multi-sensor estimator handler as depict in Figure 2. It maintains multiple sensor estimator instances and communication between them is handled locally. The estimates and ground truth values from the dataset are used for deterministic and reproducible evaluation of the estimator credibility, which is described in the following section.

The simulated UWB range measurements are modeled based on the HiPR protocol [1], with a ranging standard deviation of 0.1 m. Further, we assume to have coarse initial beliefs about each anchor's locations, e.g., by performing a calibration procedure described by Blueml et al. [14] or manual measurements.

The evaluations are performed single-threaded in MATLAB on an AMD Ryzen 7 3700X CPU with a 32 GB DDR4 RAM.

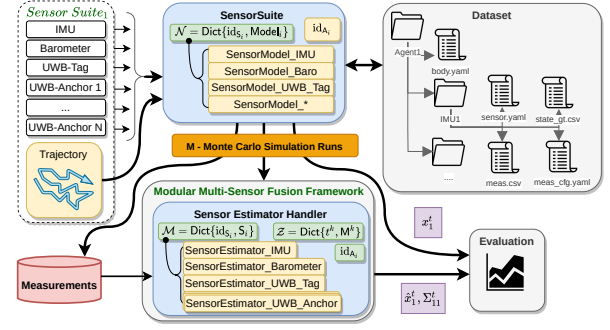


Fig. 2: Shows the block diagram of the simulation framework.

A. Estimator Consistency Metrics

To assess the filters' credibility, we compute the Root Mean Square Error (RMSE) $e_{\text{RMSE}}^k = \|\tilde{\mathbf{x}}^k\|_2$ with the error defined as $\tilde{\mathbf{x}}^k = \ominus \hat{\mathbf{x}}^k \oplus \mathbf{x}^k$ (e.g., $\tilde{\mathbf{R}} = \hat{\mathbf{R}}^{-1} \mathbf{R}$) and the single run Normalized Estimation Error Squared (NEES) $e_{\text{NEES}}^k = (\tilde{\mathbf{x}}^k)^T \Sigma^k (\tilde{\mathbf{x}}^k)$ of each state \mathbf{x} at each discrete time step t^k of the datasets. The average of a single run NEES, is denoted as $\overline{\text{NEES}} e_{\text{NEES}} = \frac{1}{K} \sum_{k=0}^K e_{\text{NEES}}^k$. Despite providing useful insights, a single run NEES does not adequately measure the filter consistency and a thorough examination is needed by evaluating the single run NEES over (multiple) M Monte Carlo simulation runs, leading to an average NEES, the ANEES at the time instance t^k

$$e_{\text{ANEES}}^k = \frac{1}{M} \sum_{i=1}^M e_{\text{NEES},i}^k. \quad (13)$$

The average RMSE, the ARMSE, of a single run is defined as $e_{\text{ARMSE}} = \frac{1}{K} \sum_{k=0}^K e_{\text{RMSE}}^k$.

Assuming a zero mean estimation error $\tilde{\mathbf{x}} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ with $\tilde{\mathbf{x}}^k = \mathbf{x}^k \ominus \hat{\mathbf{x}}^k$, the e_{ANEES} should have a chi-squared distribution χ_n^2 of dimension $n = \dim(\tilde{\mathbf{x}})$. Therefore, the NEES should be on average n . This allows to assess the credibility by defining a lower and upper boundary, r_1 , r_2 , for the observed e_{ANEES} . In general, lower values indicate pessimism (under-confidence), while higher ones optimism. The boundaries of a commonly used two-sided 95% confidence regions ($\alpha = 0.05$) is computed, defined as

$$[r_1, r_2] = \left[\frac{\chi_{nM}^2(0.5\alpha)}{M}, \frac{\chi_{nM}^2(1-0.5\alpha)}{M} \right], \quad (14)$$

with the state dimension n , M Monte Carlo runs, and the chi-squared distribution χ^2 [21]. For $M = 10$ and $n = 3$, e.g., the IMU position, the lower and upper credibility bounds for the ANEES in the 95% region are $[1.68, 4.7]$.

B. Scenario S_1

In the first scenario, S_1 , we study the tightly coupled anchor self-calibration in our modular aided inertial estimation framework. Therefore, 20 anchors are randomly distributed on a sphere with a radius of 7 m. Five anchor positions are assumed to be known and set to be fixed in the estimation framework, meaning that the corresponding sensor estimators are excluded from sensor fusion. According to a nonlinear observability analysis, only 2 anchors need be

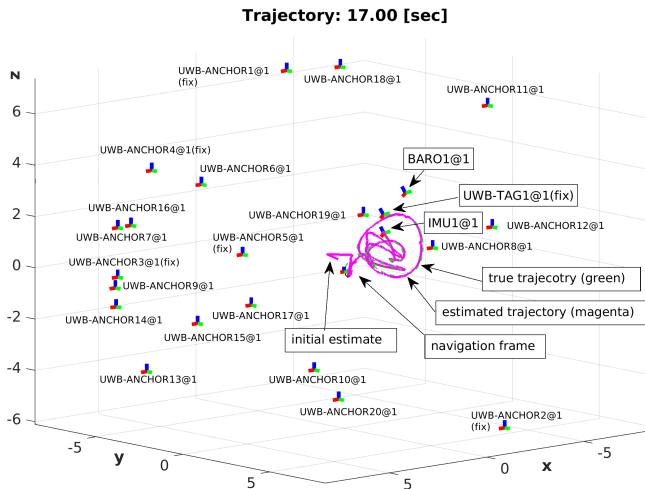


Fig. 3: Scenario S_1 : Shows the estimated trajectory at $t = 17$ s, the 15 estimated and 5 fixed anchor positions, randomly placed on a sphere with a radius of 7 m, and the estimate IMU pose with the exteroceptive sensor positions of the barometer and the fixed UWB tag. The problem is described in Section III-B.

	$\mathcal{G}_{\mathcal{P}_A}$	$\mathcal{B}_{\mathcal{P}_{\{T, \mathcal{P}\}}}$	$\mathcal{G}_{\{P, v\}_T}$	\mathcal{G}_{q_T}	τb_a	τb_w
σ^0	30 cm	10 cm	1 {m, m/s}	5 deg	0.05 m/s ²	0.05 rad/s
Δx^0	± 30 cm	± 10 cm	± 1 {m, m/s}	± 5 deg	± 0.05 m/s ²	± 0.05 rad/s

TABLE I: Scenario S_1 : Initial uncertainty and initial state offsets Δx^0 for the UWB anchor positions, IMU states, and exteroceptive sensor states. \pm emphasis that a positive or negative values is assigned randomly per element.

known in order to render the non-linear estimation problem fully observable. Since the system suffers from approximated models, linearization errors, multi-rate measurements, and a low signal-to-noise ratios, we found 5 fixed anchors to render a good compromise in terms of convergence behavior.

The agent moves along a smooth and randomly generated trajectory within the sphere for a duration of $D = 150$ s as shown in Figure 3. All states are initialized with a randomly assigned $\pm 1\sigma$ offset from the true value, as described in Table I. The IMU sample rate is 100 Hz and all UWB ranging devices are in communication range (leading to maximum complexity), while 10 % of measurements are randomly dropped. In total, three experiments regarding the anchor self-calibration are conducted. The first studies the anchor self-calibration by just obtaining tag to anchor (T2A) range measurements. In the second experiment, additionally range measurements between anchors (A2A) are used. In the third experiment, tag to anchor and readings from the barometer are used. Three different modular multi-sensor fusion strategies, MMSF-C, MMSF-DP, and MMSF-DAH are applied as already described in Section II.

1) *T2A ranging*: In this experiment, IMU and T2A measurements are fused in the modular estimator framework using three different fusion strategies. The effective T2A measurement rate for each anchor is 2 Hz due to the Round Robin scheduling performed. As shown in Table II, MMSF-DAH is the fastest approach that allows to significantly reduce the total filter execution time by a factor of 22.3

and a factor of 15.6 over MMSF-C and MMSF-DP, respectively. MMSF-DAH seems to be slightly more confident than the other two approaches and the ANEES converges slower to the desired mean value. Nonetheless, the ANEES plots of the weakly observable accelerometer bias b_a , the IMU orientation q_T , and the position of the sixth (as a representative and also random choice for space reasons) anchor p_{A_6} for all strategies shown in Figure 4, behave similarly. Summarized, the ARMSE of anchor positions reduced from an initial RMSE of 30 cm (see Table I) below to approximately 12 cm on average over the entire trajectory using either fusion strategy in Table II.

2) *T2A and A2A ranging*: In this experiment, IMU, T2A and A2A range measurements were fused using MMSF-DP and MMSF-DAH. MMSF-C was excluded due to the huge single run computation time. The effective T2A and A2A measurement rate for each device is 1 Hz due to the time scheduling performed to avoid network congestion. As shown in Table II, MMSF-DAH is 21.8 times faster as MMSF-DP. It can be clearly seen in the same table, A2A range measurements significantly reduce the estimation error of the anchor positions down to 1.6 cm. The inclusion of these measurements comes at the cost of higher complexity, leading to 5.6 times higher computation time in the case of MMSF-DAH. This could motivate to perform a two-stage approach: first perform an accurate anchor localization, then once the system is calibrated, only highly efficient T2A measurements are performed. In Figure 5, the ANEES of the accelerometer bias, the 17th anchor position and the 19th anchor position is shown. Interestingly, the A2A measurements have a positive impact on the ANEES of the accelerometer bias, and it also seems to converge faster, compared to the T2A-only experiment. The ANEES of the 17th and 19th anchor are chosen as they show both under-confidence and over-confidence, which is independent of the selected fusion strategy. This might be a position-dependent issue due to PDOP and needs further investigations. Nonetheless, the state plots in Figure 6 of the first Monte Carlo simulation run using MMSF-DAH, show an exemplary and satisfying estimation behavior of the proposed estimator.

3) *T2A and pressure*: In this experiment, we evaluate the impact of fusing pressure readings of a rigidly attached barometer tightly in the modular estimation framework using MMSF-DP and MMSF-DAH. The calibration state between the barometer and the body reference frame are assumed to be known a priori and fixed. Pressure readings are processed at a rate of 20 Hz with measurement standard deviation of $\sigma_p = 1$ Pa, translating to approximately 8.4 cm standard deviation at sea-level and is slightly beyond commercially integrated pressure sensors. Again, 10 % of pressure readings are dropped randomly. As shown in Table II, MMSF-DAH is again 21.8 times faster as MMSF-DP and fusing pressure reading tightly leads to more accurate IMU estimate in case of MMSF-DP than fusing A2A measurements. Interestingly, the pressure readings are not improving the anchor position estimates across approaches and is subject to future investigations. In Figure 5, the ANEES of the accelerometer bias,

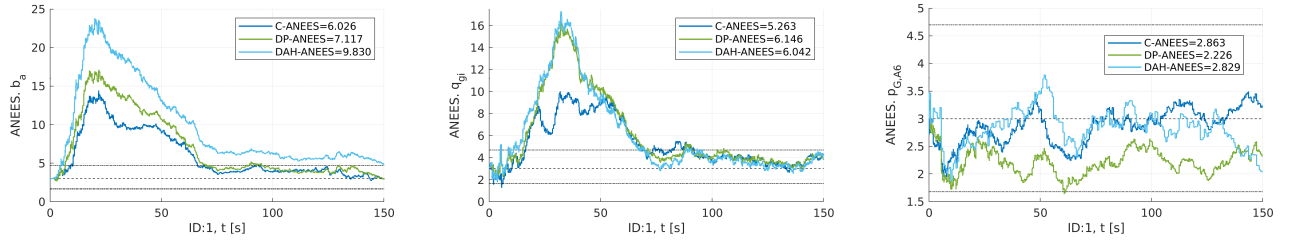


Fig. 4: Scenario S_1 : T2A ranging. Shows the ANEES, the double-sided 95 % confidence region (dotted lines), and expected ANEES value (dashed line) over 10 Monte Carlo simulation runs of the accelerometer bias \mathbf{b}_a , the IMU orientation, and the 6th UWB anchor position using MMSF-C (blue), MMSF-DP (green), and MMSF-DAH (cyan). The experiment is described in Section III-B.1.

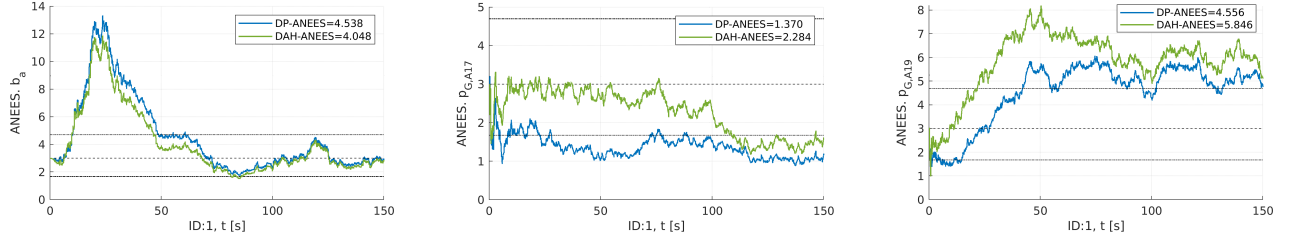


Fig. 5: Scenario S_1 : T2A and A2A ranging. Shows the ANEES, the double-sided 95 % confidence region (dotted lines), and expected ANEES value (dashed line) over 10 Monte Carlo simulation runs of the accelerometer bias \mathbf{b}_a , the 17th and 19th UWB anchor position using MMSF-DP (blue) and MMSF-DAH (green). The experiment is described in Section III-B.2.

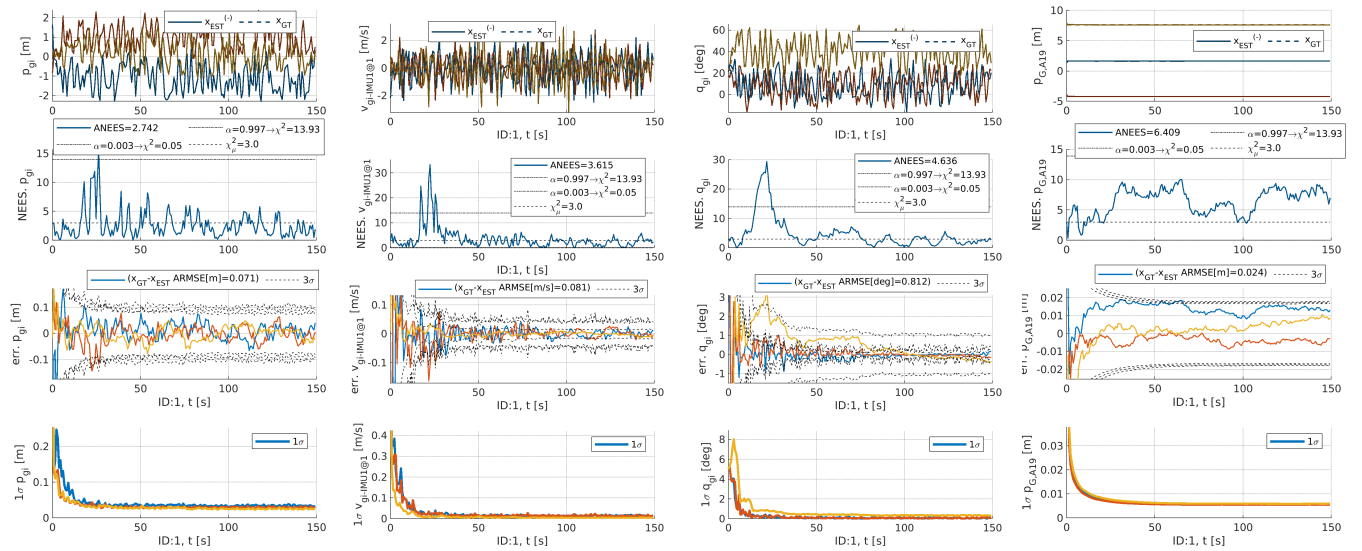


Fig. 6: Scenario S_1 : T2A and A2A ranging estimation results of the first Monte Carlo run using MMSF-DAH. The top row shows the true (dashed) and estimated values (solid), the second row shows the single run NEES with the double-sided 99.7 % confidence region (dotted lines), the third row shows the estimation error and the 3σ boundaries, and the last row shows the estimate's standard deviation. The first three columns are the estimated IMU position, velocity, and orientation. The last column is the estimated position of the 19th UWB anchor. In yellow, blue, red are for the x, y, z position, and for the roll, pitch, yaw angle, respectively. The experiment is described in Section III-B.2.

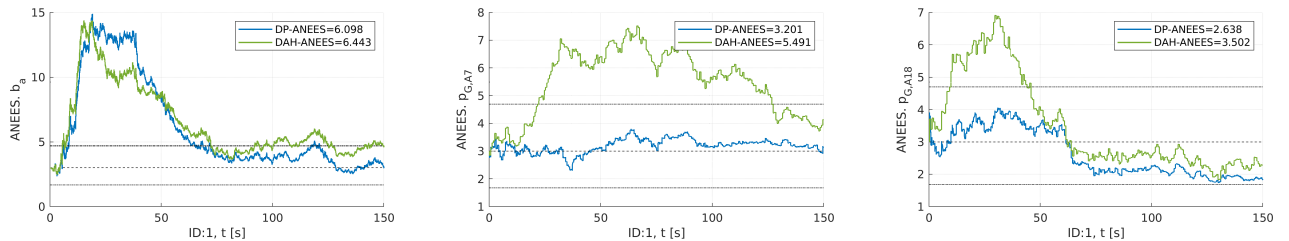


Fig. 7: Scenario S_1 : T2A ranging and pressure readings. Shows the ANEES, the double-sided 95 % confidence region (dotted lines), and expected ANEES value (dashed line) over 10 Monte Carlo simulation runs of the accelerometer bias \mathbf{b}_a , the 7th and 18th UWB anchor position using MMSF-DP (blue) and MMSF-DAH (green). The experiment is described in Section III-B.3.

10 Monte Carlo runs			[s]	[s]	[s]	[cm]		[cm/s]		[deg]		[m/s ²]		[rad/s]		[cm]	
MMSF	A2A	Baro	\bar{t}_{prop}	\bar{t}_{joint}	t_{tot}	$\bar{g}_{\mathcal{P}_T}$		$\bar{g}_{\mathcal{V}_g}$		$\bar{g}_{\mathcal{Q}_T}$		$\bar{\tau}_{\mathcal{B}_a}$		$\bar{\tau}_{\mathcal{B}_\omega}$		$\bar{g}_{\mathcal{P}_A}$	
						AR	AN	AR	AN	AR	AN	AR	AN	AR	AN	AR	AN
C	0	0	0.089	0.096	1774	4.51	2.81	7.4	3.59	0.92	5.26	0.015	6.02	0.0066	3.016	11.1	2.88
DP	0	0	0.0063	0.253	1237	4.69	2.70	7.8	3.73	1.01	6.14	0.016	7.12	0.0068	2.89	11.3	2.46
DAH	0	0	0.0039	0.0045	79.4	5.3	4.49	7.9	3.8	0.98	6.04	0.018	9.82	0.0067	3.56	11.6	3.45
DP	1	0	0.0073	0.113	9741	4.3	2.53	6.9	3.14	0.91	3.43	0.015	4.54	0.0079	4.6	1.5	3.39
DAH	1	0	0.0052	0.0044	446.7	4.4	2.52	7.1	2.98	0.91	3.18	0.015	4.04	0.0079	3.95	1.6	4.82
DP	0	1	0.0066	0.274	2075	3.85	3.76	6.1	4.36	0.81	6.17	0.012	6.1	0.0062	3.56	12.1	2.96
DAH	0	1	0.0041	0.0047	95.2	4.52	6.6	6.2	4.7	0.83	6.94	0.012	6.44	0.0066	4.22	13	5.04

TABLE II: Scenario S_1 : Shows the average ARMSE (AR) and ANEES (AN) of the estimated states as well as the average over the estimates UWB anchor positions for different fusion strategies and different sensor configurations averaged over 10 Monte Carlo simulation runs. Please note the total execution time t_{tot} is the average over a single run, while \bar{t}_{prop} and \bar{t}_{joint} are the average values for a single filter propagation or update step. Since the ARMSE calculation includes the entire trajectory, including the initialization and the state convergence phase, it is on average higher than the error after convergence. Best values in bold and problem is described in Section III-B.

the 7th and 18th UWB anchor are depicted, indicating again that MMSF-DAH is more optimistic than MMSF-DP and that the ANEES converges slower, while all estimates tend to converge towards the defined confidence region.

C. Scenario S_2

For evaluating the estimator credibility and the computation time of individual sensor estimators in a realistic scenario, we use two *Machine Hall* sequences (MH_04 and MH_05) of the EuRoC dataset [26]. The simulation runs for 78s and 92s, which is the flight time of the *Machine Hall* sequences between take-off and landing. Each UAV is equipped with an IMU, barometer, and an UWB transceiver (tag), for both communication and pair-wise ranging between other UWB modules in communication range. Figure 1 depicts the spatial frame constellation. Twenty-five stationary UWB transceiver (anchors) are assumed to be deployed to cover the area of interest as shown in Figure 8 with a communication range of 4 m. The communication range is on purpose short, to justify the deployment of 25 anchors and to challenge the estimation problem: anchors are revisited again, meaning that they are correlated and need to be considered properly, and some anchors have no direct link to (fixed) reference anchors. Five anchors close to the take-off position of the UAV are fixed (constant) to define the global coordinate reference frame.

The same initial values and parameters as in Section III-B are used, with the difference that the initial uncertainty of the anchor position was lowered to $\sigma_{\mathcal{P}_A} = 0.1$ m and the effective UWB ranging rate between devices in range is 8 Hz (due to the sparse configuration, fewer devices are in communication range, which allows a higher net rate for individuals) and the IMU rate in the dataset is 200 Hz. All UWB nodes are inserted artificially into the real-world dataset.

Figure 8 shows the estimation performance using MMSF-DP and MMSF-DAH. As confirmed in Table III, hardly any difference in the estimated IMU states are noticeable, while operating on the full state vector (with 75 elements) in case of MMSF-DP in the filter update steps causes tremendous computation efforts over treating them isolated in case of

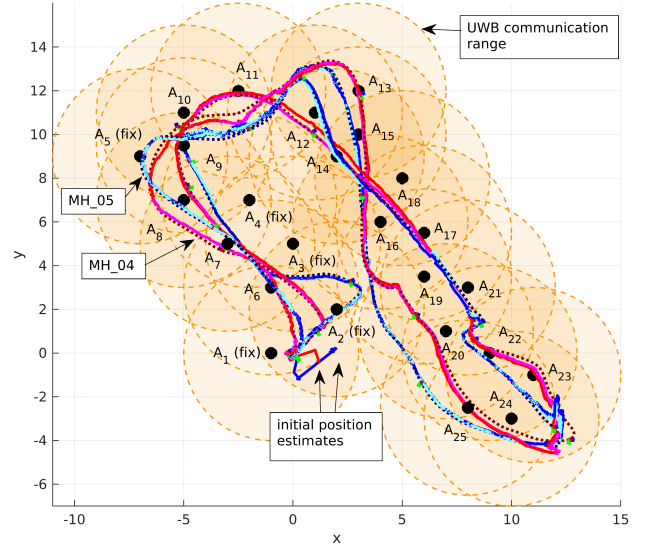


Fig. 8: Scenario S_2 : Top view on the anchor placement (bullets) and the Machine Hall 4 and 5 trajectories of the *Machine Hall* sequences of the EuRoC dataset [26] in dashed dark red and dashed dark blue, respectively. In red and blue are the estimated position using MMSF-DAH, while in magenta and cyan using MMSF-DP. The filled orange circles show the communication range of 4 m per UWB anchor. Fixed anchors at the starting position are assumed to be constant and known, to define the navigation reference frame.

MMSF-DAH. This processing speedup of almost 26.5 comes at the cost of a slightly degraded and slower converging anchor self-calibration, as discussed in Section III-B.

IV. CONCLUSION AND FUTURE WORK

Considering inter-sensor observations not involving the core navigation states in a Modular Multi-Sensor Fusion (MMSF) framework is a new paradigm originating from Collaborative State Estimation (CSE). In our evaluations, we have shown that considering these as tightly coupled range observations between ultra-wideband (UWB) anchors can significantly improve the estimates of both, the navigation states and estimated anchor positions. In such a formulation, the computational effort increases significantly using centralized-equivalent estimators. With the presented sensor extensions to our recently proposed Modular Multi-

10 Monte Carlo runs		[s]	[s]	[s]	[cm]		[cm/s]		[deg]		[m/s ²]		[rad/s]		[cm]	
MMSF	EuRoC	\bar{t}_{prop}	\bar{t}_{joint}	t_{tot}	$\bar{g}^G \mathbf{p}_I$		$\bar{g}^G \mathbf{v}_g$		$\bar{g}^G \mathbf{a}_I$		$\bar{g}^G \mathbf{b}_a$		$\bar{g}^G \mathbf{b}_w$		$\bar{g}^G \mathbf{p}_A$	
					AR	AN	AR	AN	AR	AN	AR	AN	AR	AN	AR	AN
DP	MH.04	0.0071	0.209	2337.8	9.82	3.28	9.6	3.97	1.52	2.96	0.03	2.91	0.003	2.1	5.9	4.77
DAH	MH.04	0.0036	0.0029	88.7	10.7	3.81	9.8	3.92	1.58	2.87	0.03	2.72	0.003	2	6.6	4.7
DP	MH.05	0.0072	0.204	3087.9	10.8	3.86	9.3	4.63	1.72	3.38	0.029	2.2	0.0028	2.62	5.65	5.78
DAH	MH.05	0.0038	0.0032	116.6	12.56	4.3	9.77	4.87	1.73	3.77	0.03	3.37	0.0029	4.43	6.55	5.66

TABLE III: Scenario S₂: Shows the average ARMSE (AR) and ANEES (AN) of the estimated states as well as the average over the estimates UWB anchor positions for different fusion strategies and different sensor configurations averaged over 10 Monte Carlo simulation runs. Please note the total execution time t_{tot} is the average over a single run, while \bar{t}_{prop} and \bar{t}_{joint} are the average values for a single filter propagation or update step. Best values in bold and problem is described in Section III-C.

Sensor Fusion Decoupled Approximated History (MMSF-DAH) approach, we show the feasibility of processing inter-sensor observation as isolated joint updates between sensor estimators without requiring estimates of other sensors and, thus, resulting in a significant speedup. This breaks the computational barrier such that meshed (and generally ill scaling) inter-sensor observations can be fused to aided inertial navigation estimators while maintaining consistency in the navigation states.

UWB aided inertial motion estimation with simultaneous online UWB anchor position initialization (and extension by fly-by), and continuous self-calibration remains a challenging task. In future work, we will also investigate in increasing the robustness in case of NLOS conditions (that lead to unmodeled offsets in range measurements), which requires consistent estimates for a statistical outlier rejection.

REFERENCES

- [1] D. Neuhold, C. Bettstetter, and A. F. Molisch, “HiPR: High-Precision UWB Ranging for Sensor Networks,” in *Proceedings of the 22nd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, ser. MSWIM ’19. New York, NY, USA: Association for Computing Machinery, Nov. 2019, pp. 103–107.
- [2] Z. Sahinoglu, S. Gezici, and I. Güvenc, *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols*. Cambridge University Press, 2008.
- [3] A. Ledergerber, M. Hamer, and R. D’Andrea, “A robot self-localization system using one-way ultra-wideband communication,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept. 2015, pp. 3131–3137.
- [4] J. D. Hol, F. Dijkstra, H. Luinge, and T. B. Schon, “Tightly coupled UWB/IMU pose estimation,” in *2009 IEEE International Conference on Ultra-Wideband*, Sept. 2009, pp. 688–692.
- [5] J. D. Hol, T. B. Schön, and F. Gustafsson, “Ultra-wideband calibration for indoor positioning,” in *2010 IEEE International Conference on Ultra-Wideband*, vol. 2, Sept. 2010, pp. 1–4.
- [6] K. Batstone, M. Oskarsson, and K. Åström, “Towards real-time time-of-arrival self-calibration using ultra-wideband anchors,” in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sept. 2017, pp. 1–8.
- [7] Y. Song, M. Guan, W. P. Tay, C. L. Law, and C. Wen, “UWB/LiDAR Fusion For Cooperative Range-Only SLAM,” in *2019 International Conference on Robotics and Automation (ICRA)*, May 2019, pp. 6568–6574.
- [8] Q. Shi, S. Zhao, X. Cui, M. Lu, and M. Jia, “Anchor self-localization algorithm based on UWB ranging and inertial measurements,” *Tsinghua Science and Technology*, vol. 24, no. 6, pp. 728–737, Dec. 2019.
- [9] T. H. Nguyen, T.-M. Nguyen, and L. Xie, “Range-Focused Fusion of Camera-IMU-UWB for Accurate and Drift-Reduced Localization,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1678–1685, Apr. 2021.
- [10] A. Goudar and A. P. Schoellig, “Online Spatio-temporal Calibration of Tightly-coupled Ultrawideband-aided Inertial Localization,” in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept. 2021, pp. 1161–1168.
- [11] M. W. Mueller, M. Hamer, and R. D’Andrea, “Fusing ultra-wideband range measurements with accelerometers and rate gyroscopes for quadcopter state estimation,” in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015, pp. 1730–1736.
- [12] K. Hausman, S. Weiss, R. Brockers, L. Matthies, and G. S. Sukhatme, “Self-calibrating multi-sensor fusion with probabilistic measurement validation for seamless sensor switching on a UAV,” in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, May 2016, pp. 4289–4296.
- [13] M. Hamer and R. D’Andrea, “Self-Calibrating Ultra-Wideband Network Supporting Multi-Robot Localization,” *IEEE Access*, vol. 6, pp. 22 292–22 304, 2018.
- [14] J. Bluemel, A. Fornasier, and S. Weiss, “Bias compensated uwb anchor initialization using information-theoretic supported triangulation points,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 5490–5496.
- [15] K. Guo, Z. Qiu, W. Meng, L. Xie, and R. Teo, “Ultra-wideband based cooperative relative localization algorithm and experiments for multiple unmanned aerial vehicles in GPS denied environments,” *International Journal of Micro Air Vehicles*, vol. 9, no. 3, pp. 169–186, Sept. 2017.
- [16] H. Xu, L. Wang, Y. Zhang, K. Qiu, and S. Shen, “Decentralized Visual-Inertial-UWB Fusion for Relative State Estimation of Aerial Swarm,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, May 2020, pp. 8776–8782.
- [17] K. Guo, X. Li, and L. Xie, “Ultra-Wideband and Odometry-Based Cooperative Relative Localization With Application to Multi-UAV Formation Control,” *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2590–2603, June 2020.
- [18] T. Ziegler, M. Karrer, P. Schmuck, and M. Chli, “Distributed Formation Estimation Via Pairwise Distance Measurements,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3017–3024, Apr. 2021.
- [19] M. Shalaby, C. C. Cossette, J. R. Forbes, and J. Le Ny, “Relative Position Estimation in Multi-Agent Systems Using Attitude-Coupled Range Measurements,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4955–4961, July 2021.
- [20] P. Maybeck, *Stochastic Models, Estimation and Control*, ser. Mathematics in science and engineering. Academic Press, 1979.
- [21] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*, ser. Electronic and electrical engineering. Wiley, 2001.
- [22] R. Jung and S. Weiss, “Modular Multi-Sensor Fusion: A Collaborative State Estimation Perspective,” *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6891–6898, Oct. 2021.
- [23] —, “Scalable Recursive Distributed Collaborative State Estimation for Aided Inertial Navigation,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, May 2021, pp. 1896–1902.
- [24] S. Weiss and R. Siegwart, “Real-time metric state estimation for modular vision-inertial systems,” in *2011 IEEE International Conference on Robotics and Automation (ICRA)*, May 2011, pp. 4531–4537.
- [25] M. Bevermeier, O. Walter, S. Peschke, and R. Haeb-Umbach, “Barometric height estimation combined with micro-matching in a loosely-coupled Kalman-filter,” in *Navigation and Communication 2010 7th Workshop on Positioning*, Mar. 2010, pp. 128–134.
- [26] 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*, vol. 35, no. 10, pp. 1157–1163, Sept. 2016.