2023 IEEE International Conference on Robotics and Automation (ICRA 2023)
May 29 - June 2, 2023. London, UK

Distributed Initialization for Visual-Inertial-Ranging Odometry with Position-Unknown UWB Network

Shenhan Jia, Rong Xiong and Yue Wang

***Abstract*— In recent years, the visual-inertial-ranging (VIR) state estimator with a position-unknown UWB network has become popular. However, most existing VIR methods leverage centralized algorithms to initialize the UWB anchors, which are challenging to be applied to massive UWB networks. In this paper, we propose a distributed initialization method for consistent visual-inertial-ranging odometry with a position-unknown UWB network (DC-VIRO). For the position-unknown UWB anchors, we solve a Robot-aided Distributed Localization (RaDL) to initialize their positions. For robot state estimation, we fuse the ranging measurements of initialized anchors and visual-inertial measurements in a consistent filter. The RaDL is formulated as a consensus-based optimization problem and solved by the Distributed Alternating Direction Method of Multipliers (D-ADMM) algorithm. To identify the unobservable conditions, we propose a self-contained Fisher Information Matrix (FIM) based criterion which can be evaluated by each anchor directly with locally-preserved ranging measurements. We use Covariance Intersection (CI) to estimate the covariance of initialized anchors' positions for consistent data fusion. The proposed DC-VIRO is validated in both simulation and real-world experiments.**

I. Introduction

State estimation is one of the most fundamental modules in many applications, such as robotic navigation, VR/AR, and autonomous driving. In the past ten years, visual-inertial odometry (VIO) has achieved significant progress [1] [2] [3]. However, VIO methods suffer from localization drift over long trajectories due to the lack of global information [4]. In recent years, more and more researchers propose methods that incorporate ultra-wideband (UWB) measurements into the VIO to reduce localization drift [5] [6] [7]. A typical scenario of UWB-aided visual-inertial state estimation can be seen in Fig. 1. Besides the UWB node mounted on the robot, there is also a UWB network containing an arbitrary number of anchors. Distance measurements and communication between two UWB sensors are possible only if their distance is within a threshold.

Traditional works [8] [7] [9] require the UWB anchors' positions to be pre-calibrated before starting up the estimator, that is, the anchors' global positions are known to the robot. However, in physically harsh or difficult-to-reach environments, UWB sensors are always dropped from air-crafts or via artillery launch [10] [11], whose global positions are unknown. In these conditions, the problem can be divided into two stages: 1. a robot-aided UWB sensor network localization problem; 2. a robot state estimation problem with onboard visual-inertial sensors and external initialized UWB anchors.

This work was supported by the National Key R&D Program of China under Grant 2021ZD0114500.

All authors are with the State Key Laboratory of Industrial Control and Technology, Zhejiang University, Hangzhou, P.R. China. Yue Wang is the corresponding author wangyue@iipc.zju.edu.cn.

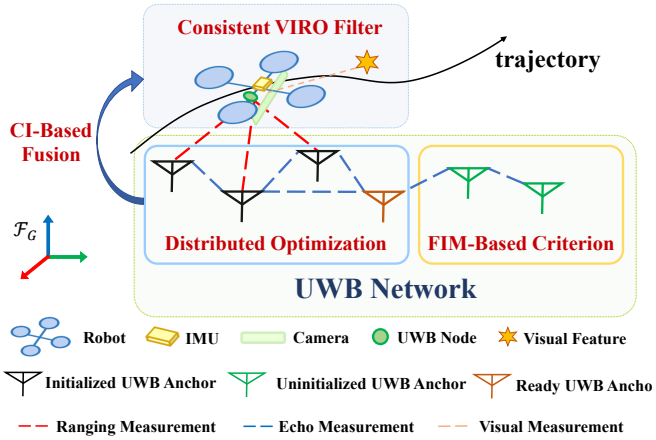


Fig. 1: A system setting illustration of our proposed DC-VIRO system. The robot fuses visual-inertial-ranging measurements within an EKF-based filter. The UWB network can be seen as a peer-to-peer cluster, where UWB anchors estimate their positions with distributed non-linear optimization.

In the first stage, we localize a UWB network with the assistance of a mobile robot. Previous works [5] [12] [13] propose localizing the UWB network by solving a centralized non-linear optimization problem. However, when there are numerous UWB anchors, centralized methods are prone to severe communication congestion throughout the network and high computational complexity at the central node [14]. Motivated by the distributed methods for localizing static sensor networks [15] [16] [17], we propose a Robot-aided Distributed Localization (RaDL) algorithm to localize UWB anchors, where no sensor measurement is transmitted and the computing loads are distributed on the anchor's onboard computers. There are two main challenges in RaDL: 1. solving the optimization problem in a distributed manner; 2. identifying the unobservable conditions to avoid divergence of the solver. For the first challenge, we reformulate the centralized problem through the consensus formulation ([18]) and solve it with a popular dual approach called distributed ADMM [19] [20]. For the second challenge, the sensor position's observability is typically analyzed by the Fisher Information Matrix (FIM) [21]. However, existing methods [22] [23] use the anchor's position to evaluate the FIM, which requires solving the computing-intensive optimization before the evaluation of FIM. By exploiting the geometric interpretation, we propose an incremental method to evaluate the determinant of FIM *directly* with the anchor's local measurements, resulting in an online FIM-based criterion to

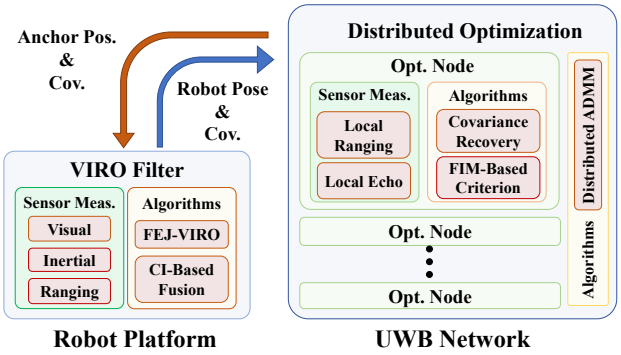


Fig. 2: Illustration of modules and data flows of the proposed DC-VIRO.

start up the optimization.

In the second stage, we extend the RaDL algorithm to a complete visual-inertial-ranging odometry estimator with a position-unknown UWB network based on our previous work [5]. Since the history poses corresponding to the ranging measurements are preserved in the long window of the state [5], the state vector and its covariance expand with the increasing number of UWB anchors, resulting in challenges in applying to massive UWB networks. In this work, we recover the covariance of the optimized anchors with covariance intersection (CI) [24] [25], which is a useful technique for consistent fusion and covariance recovery [26]. We broadcast the poses and small marginal covariance of the robot keyframes to the UWB network, then the upper bound of the full covariance is computed with CI. As a result, each anchor can recover its covariance after optimization, then transmit the position and covariance to the robot for consistent state augmentation. After an anchor i is augmented in the robot state, the robot updates the state with distance measurements between the robot and anchor i acquired by the robot.

In summary, the contributions of this paper are listed as follows:

- We propose DC-VIRO, a distributed initialization method for consistent visual-inertial-ranging odometry with a position-unknown UWB network.
- We reformulate the UWB initialization problem within the consensus optimization framework and solve it with the D-ADMM algorithm. We also propose an FIM-based criterion to identify unobservable conditions, which is self-contain and only requires local ranging measurements.
- We recover the anchor position's covariance efficiently with the CI technique. The recovered covariance and the optimized anchor position are augmented in the robot state, resulting in an accurate and consistent VIRO filter.

II. System Overview

A. Notations

There are three coordinate frames in this paper, which are the *global frame* $\mathcal{F}_G$, the *IMU frame* $\mathcal{F}_I$ and the camera frame $\mathcal{F}_C$. We refer the readers to [5] for an illustration of these frames and other notations.

B. RaLD Method

The UWB network is designed as a peer-to-peer structure, where UWB anchors act as nodes (*Opt. Node* in Fig. 2) in the cluster. Every node in the UWB network estimates the determinant of the Fisher Information Matrix (FIM) with its locally-preserved ranging measurements (*Local Ranging* in Fig. 2), then decides whether it is ready to initialize (*FIM-Based Criterion* in Fig. 2). We use a centralized Levenberg-Marquard algorithm for single-anchor initialization.

Algorithm 1: RaDL on Anchor i

input : Distance Measurements $\{\mathcal{R}\}$;
Robot Pose with Cov. $\{\mathbf{p}_r, \mathbf{R}_r, \mathbf{P}_{mr}\}$;
Ready List from Neighbors $\tilde{\mathcal{L}}$.
output: Anchor Pos. with Cov. $\{\mathbf{p}_a, \mathbf{P}_a, \mathbf{P}_{xa}\}$.

1 $\Upsilon \leftarrow \emptyset$; $\mathcal{L} \leftarrow \emptyset$;
2 FIMReady $\leftarrow$ False;
3 **while** *True* **do**
4 $\quad \Upsilon = \Upsilon \cup \{\mathcal{R}, \mathbf{p}_r, \mathbf{R}_r, \mathbf{P}_{mr}\}$;
5 $\quad$ **if** $\mathcal{L} \neq \tilde{\mathcal{L}}$ **then**
6 $\quad\quad \mathcal{L} \leftarrow \mathcal{L} \cup \tilde{\mathcal{L}}$;
7 $\quad\quad$ *Broadcast* $\mathcal{L}$;
8 $\quad$ **end**
9 $\quad$ **if** FIMReady *is* True **then**
10 $\quad\quad$ **continue**;
11 $\quad$ **end**
12 $\quad$ FIMReady $\leftarrow$ `FIM_Criterion(`Υ`);`
13 $\quad$ **if** FIMReady *is* True **then**
14 $\quad\quad \mathcal{L} \leftarrow \mathcal{L} \cup \{i\}$; *Broadcast* $\mathcal{L}$;
15 $\quad\quad$ *Wait lock*; $\mathbf{p}_a \leftarrow$ `D_ADMM(`Υ*; $\mathcal{L}$, lock*`);`
16 $\quad\quad \mathbf{P}_a, \mathbf{P}_{xa} \leftarrow$ `Cov_Rec(`$\mathbf{p}_a$, Υ`);`
17 $\quad\quad$ *Send* $\{\mathbf{p}_a, \mathbf{P}_a, \mathbf{P}_{xa}\}$ *to the robot*;
18 $\quad$ **end**
19 **end**

Algorithm 1 presents the procedures runs on each UWB *Opt. Node*. Note that the *broadcast* sends messages to its *one-hop* neighbors. The anchor i collects distance measurements and robot pose with marginal covariance in Υ (line 4). Every anchor also preserves a *ready list* $\mathcal{L}$ containing IDs of the anchors that have met the FIM-based Criterion. Lines $5 \sim 8$ ensure that the $\mathcal{L}$ of anchor i is consensus with its reachable neighbors after several loops. We then consider the lines $9 \sim 15$. Once the anchor meets the FIM-based criterion, it updates the ready list (line 14). The *wait lock* means if there are anchors in $\mathcal{L}$ running the *D_ADMM*, it would wait for it. Lines $16 \sim 17$ represent the anchor i recover the covariance of the optimized position with its local measurements. The covariance recovery is based on the CI technique. Now we know that if an anchor meets the FIM criterion, it would be initialized. So lines $9 \sim 11$ represent that the anchor only collects measurements and updates the ready list (before line 9) if it has been initialized.

C. VIRO Estimator

As shown in Fig. 2, the robot platform fuses visual-inertial-ranging measurements within a Multi-State Constraint Kalman Filter (MSCKF) framework(*FEJ-VIRO* in Fig. 2). Algorithm 2 presents the procedures running on the robot platform, which can be viewed as two parts: lines $2 \sim 9$ robot state estimation, while lines $10 \sim 13$ augment the initialized UWB anchors into the state with CI. We discard a ranging measurement at the robot platform if the corresponding UWB anchor has not been augmented to the robot state (lines $3 \sim 7$).

Algorithm 2: VIRO Estimator

input : VIR Measurements $\{\mathcal{V}, \mathcal{I}, \mathcal{R}\}$;
Anchor Pos. and Cov. $\{\mathbf{p}_a, \mathbf{P}_a\}$.
output: Robot States $\mathbf{x}_r$ with Covariance $\mathbf{P}_r$.

1 **while** *True* **do**
2 *Read sensor data* $\mathcal{S} = \{\mathcal{V}, \mathcal{I}, \mathcal{R}\}$;
3 **for** $d_i \in \mathcal{R}$ **do**
4 **if** $\mathbf{p}_{a_i} \notin \mathbf{x}_r$ **then**
5 $\mathcal{R} \leftarrow \mathcal{R} \setminus \{d_i\}$
6 **end**
7 **end**
8 $\mathbf{x}_r,\ \mathbf{P}_r \leftarrow$ `FEJ_VIRO` $(\mathbf{x}_r, \mathbf{P}_r, \mathcal{S})$;
9 *Broadcast robot pose and marginal covariance*;
10 *Read from UWB network:* $\mathcal{U} = \{\mathbf{p}_a, \mathbf{P}_a\}$;
11 **if** $\mathcal{U} \neq \emptyset$ **then**
12 $\mathbf{x}_r,\ \mathbf{P}_r \leftarrow$ `CI_Fusion` $(\mathbf{x}_r, \mathbf{P}_r, \mathcal{U})$;
13 **end**
14 **end**

III. Distributed UWB Initialization

In this section, we describe our distributed robot-aided initialization for position-unknown UWB anchors. We first present the FIM-based criterion, which decides when to initialize the current anchor. Then we reformulate the optimization problem in a consensus optimization framework and solve it with D-ADMM.

A. FIM-Based Criterion

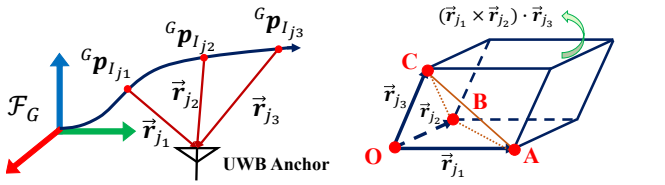
The Fisher information is a way of measuring the amount of information that a set of measurements carry about the unknown state.

Assume there are n ranging measurements preserved in the current UWB anchor, denoted as:

$$\mathbf{d} = \begin{bmatrix} d_1 & d_2 & \cdots & d_n \end{bmatrix}^\top \tag{1}$$

where $d_i,\ i = 1, \cdots, n$ is the distance measurements accrued at time-step t_i. We fix the robot's poses and only estimate the UWB anchors' positions in the initialization module, and each ranging measurement can be expressed with the anchor's position ${}^G\mathbf{p}_a$:

$$\begin{aligned} d_i &= h_i({}^G\mathbf{p}_a) + n_r \\ &= \|{}^G\mathbf{p}_{r,i} - {}^G\mathbf{p}_a\| + d_{bias} + n_r \end{aligned} \tag{2}$$



(a) (b) $|\mathbf{V}_{OABC}| = \frac{1}{6}(\mathbf{r}_{j1} \times \mathbf{r}_{j2}) \cdot \mathbf{r}_{j3}$

Fig. 3: Illustration of the vectors $\{\mathbf{r}_{j1}, \mathbf{r}_{j2}\ \mathbf{r}_{j3}\}$ in (7) and the parallelepiped defined by them. The $(\mathbf{r}_{j1} \times \mathbf{r}_{j2}) \cdot \mathbf{r}_{j3}$ equals to the parallelepiped's volume, which is six times of the tetrahedron's volume $|\mathbf{V}_{OABC}|$.

where ${}^G\mathbf{p}_{r,i} = {}^G\hat{\mathbf{p}}_{I,i} + {}^{I,i}_{G}\hat{\mathbf{R}}^\top {}^I\mathbf{p}_r$; $\{{}^G\hat{\mathbf{p}}_{I,i},\ {}^{I,i}_{G}\hat{\mathbf{R}}\}$ is the robot pose at time-step t_i, which can be received from the communication network; n_r is the zero-mean Gaussian noise; and d_{bias} is the constant ranging bias.

We further define the noise-free distance vector as:

$$\mathbf{h}({}^G\mathbf{p}_a) = \begin{bmatrix} h_1 & h_2 & \cdots & h_n \end{bmatrix}^\top \tag{3}$$

where $h_i,\ i = 1, \cdots, n$ is short for $h_i({}^G\mathbf{p}_a) = \|{}^G\mathbf{p}_{r,i} - {}^G\mathbf{p}_a\| + d_{bias},\ i = 1, \cdots, n$. Under the assumption of zero-mean Gaussian noise n_r, the FIM, denoted by $\mathbf{F}$, is given by the following form:

$$\mathbf{F} = \left[\frac{\partial \tilde{\mathbf{h}}({}^G\mathbf{p}_a)}{\partial\, {}^G\tilde{\mathbf{p}}_a}\right]^\top \mathbf{\Sigma}_r^{-1} \left[\frac{\partial \tilde{\mathbf{h}}({}^G\mathbf{p}_a)}{\partial\, {}^G\tilde{\mathbf{p}}_a}\right] \tag{4}$$

where $\mathbf{\Sigma}_r^{-1} = \frac{1}{\sigma_r^2}\mathbf{I}_{n\times n}$; σ_r is the noise density of ranging measurements. By taking the derivative of (2), we have

$$\mathbf{H}_a \triangleq \frac{\partial \tilde{\mathbf{h}}({}^G\mathbf{p}_a)}{\partial\, {}^G\tilde{\mathbf{p}}_a} = \begin{bmatrix} \frac{1}{\hat{d}_1}({}^G\mathbf{p}_a - {}^G\mathbf{p}_{r,1})^\top \\ \frac{1}{\hat{d}_2}({}^G\mathbf{p}_a - {}^G\mathbf{p}_{r,2})^\top \\ \vdots \\ \frac{1}{\hat{d}_n}({}^G\mathbf{p}_a - {}^G\mathbf{p}_{r,n})^\top \end{bmatrix} \tag{5}$$

where $\hat{d}_i = d_i - d_{bias} (i = 1, \cdots n)$ are the unbiased ranging measurements. As a result, we have $\mathbf{F} = \frac{1}{\sigma_r^2}\mathbf{H}_a^\top \mathbf{H}_a$.

As shown in (5), if we evaluate $\det(\mathbf{F})$ with $\mathbf{H}_a$, the anchor's position ${}^G\mathbf{p}_a$ should be estimated, which requires solving a computing-intensive optimization. So we exploit the geometric interpretation by applying the Cauchy-Binet formula to $\det(\mathbf{F})$, and it has the following form:

$$\det(F) = \frac{1}{\sigma_r^2} \sum_{1 \le j_1 < j_2 < j_3 \le n} (\det(\begin{bmatrix} \mathbf{u}_{j_1} \\ \mathbf{u}_{j_2} \\ \mathbf{u}_{j_3} \end{bmatrix}))^2 \tag{6}$$

where $\mathbf{u}_i = \frac{1}{\hat{d}_i}({}^G\mathbf{p}_a - {}^G\mathbf{p}_{r,i})^\top,\ i \in \{j_1, j_2, j_3\}$ is the i-th row of $\mathbf{H}_a$. By defining $\mathbf{r}_i = \hat{d}_i\, \mathbf{u}_i = {}^G\mathbf{p}_a - {}^G\mathbf{p}_{r,i}$, we have

$$\det(\begin{bmatrix} \mathbf{u}_{j_1} \\ \mathbf{u}_{j_2} \\ \mathbf{u}_{j_3} \end{bmatrix}) = \frac{1}{\hat{d}_{j_1}\hat{d}_{j_2}\hat{d}_{j_3}} (\mathbf{r}_{j_1} \times \mathbf{r}_{j_2}) \cdot \mathbf{r}_{j_3} \tag{7}$$

where $(\mathbf{r}_{j_1} \times \mathbf{r}_{j_2}) \cdot \mathbf{r}_{j_3}$ equals to the volume of the parallelepiped shown in Fig. 3. Then we compute the tetrahedron's volume $|\mathbf{V}_{OABC}|$ presented in Fig. 3(b), and $(\mathbf{r}_{j_1} \times \mathbf{r}_{j_2}) \cdot \mathbf{r}_{j_3}$ is six times of it. Reversing the directions of $\{\mathbf{r}_{j_1}, \mathbf{r}_{j_2}, \mathbf{r}_{j_3}\}$ in Fig. 3(b), the three points *A, B, C* correspond to $\{{}^G\mathbf{p}_{I_{j_1}}, {}^G\mathbf{p}_{I_{j_2}}, {}^G\mathbf{p}_{I_{j_3}}\}$ in Fig. 3(a). So we can

compute $|\mathbf{V}_{OABC}|$ with the Heron-type formula and Cayley-Menger determinant:

$$|\mathbf{V}_{OABC}| = \sqrt{\frac{1}{288}\cdot\begin{vmatrix} 0 & U^2 & V^2 & W^2 & 1 \\ U^2 & 0 & w^2 & v^2 & 1 \\ V^2 & w^2 & 0 & u^2 & 1 \\ W^2 & v & 2 & u^2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix}} \tag{8}$$

$$U = |\mathbf{r}_{j_1}| \approx d_{j_1}, \quad u = \|{}^G\mathbf{p}_{I_{j_2}} - {}^G\mathbf{p}_{I_{j_3}}\|$$
$$V = |\mathbf{r}_{j_2}| \approx d_{j_2}, \quad v = \|{}^G\mathbf{p}_{I_{j_1}} - {}^G\mathbf{p}_{I_{j_3}}\|$$
$$W = |\mathbf{r}_{r_3}| \approx d_{j_3}, \quad w = \|{}^G\mathbf{p}_{I_{j_1}} - {}^G\mathbf{p}_{I_{j_2}}\|$$

Note that the *real* $|\mathbf{r}_i|$ ($i \in \{j_1, j_2, j_3\}$) does not equal to the ranging measurement d_i ($i \in \{j_1, j_2, j_3\}$) because of the sensor noise, but we can approximate the volume with these measurements.

Considering that $\det([\mathbf{u}_{j_1}\ \mathbf{u}_{j_2}\ \mathbf{u}_{j_3}]^\top) = 6\,|\mathbf{V}_{OABC}|$ as shown in Fig. 3(b), we can compute $\det(\mathbf{F})$ with (6). Furthermore, the $\det(\mathbf{F})$ can be updated incrementally with the newly received ranging measurement. Assume the anchor received a new ranging measurement d_k at time-step t_k, and the $\det(\mathbf{F}_{k-1})$ is computed with Υ_{k-1}, which contains all the ranging measurements collected until t_{k-1}. We can combine d_k with ranging pairs $\{d_i, d_j\} \in \Upsilon_{k-1}$ and the corresponding robot poses to compute $\det([\mathbf{u}_k\ \mathbf{u}_j\ \mathbf{u}_i]^\top)^2$ with (8). Then we add the results to $\det(\mathbf{F}_{k-1})$ to get the current FIM estimate $\det(\mathbf{F}_k)$. Since $\det(\mathbf{F})$ quantifies the information carried by the existing n ranging measurements, we can set a threshold τ_F and initialize the UWB anchor's position when $\det(\mathbf{F}) > \tau_F$.

B. Distributed Optimization

We initialize the UWB anchors' positions by solving a non-linear optimization problem, which is reformulated within the consensus-based optimization framework in this subsection. The initial values of the solver is given by [27] as in [5].

We assume that a UWB anchor meets the FIM-based criterion at time-step t_k (see the *Ready UWB Anchor* in Fig. 1), and there are $n-1$ initialized UWB anchors (see the *Initialized UWB Anchor* in Fig. 1). For these n UWB anchors, we have $\det(\mathbf{F}_i) > \tau_F$ ($i = 1, 2, \cdots, n$). We define the state vector of the i-th UWB anchor to be:

$$\mathbf{x}_i^c = \begin{bmatrix} {}^G\mathbf{p}_{a_1}^\top & {}^G\mathbf{p}_{a_2}^\top & \cdots & {}^G\mathbf{p}_{a_n}^\top \end{bmatrix} \tag{9}$$

where ${}^G\mathbf{p}_{a_i}^\top$ ($i = 1, \cdots, n$) is the global position of the i-th UWB anchor. The consensus optimization problem is defined as:

$$\min_{\mathbf{x}_i^c} \sum_{i=1}^{n} f_i(\mathbf{x}_i^c)$$
$$s.t. \quad \mathbf{x}_i^c = \mathbf{x}_j^c, \quad \forall i, j = 1, \cdots, n \tag{10}$$

where $f_i(\mathbf{x}_i^c)$ is the sub-problem which associates with the i-th UWB anchor's local ranging and echo measurements (see *Opt. Node* in Fig. 2). And $f_i(\mathbf{x}_i^c)$ is defined as:

$$f_i(\mathbf{x}_i^c) = \sum_{d_{r,k}} \frac{(d_{r,k} - d_{bias} - \|{}^G\mathbf{p}_{a_i} - {}^G\mathbf{p}_{r,k}\|)^2}{\sigma_r^2} + \sum_{j\in\mathcal{S}(i)}\sum_{d_{e_{ij}}} \frac{(d_{e_{ij}} - d_{bias} - \|{}^G\mathbf{p}_{a_i} - {}^G\mathbf{p}_{a_j}\|)^2}{\sigma_e^2} \tag{11}$$

where $d_{r,k}$ is the distance measured by the i-th UWB anchor at time-step t_k; ${}^G\mathbf{p}_{r,k} = {}^G\mathbf{p}_{I,k} + {}^{I,k}_{G}\mathbf{R}^\top {}^I\mathbf{p}_r$ is the UWB node's position at time-step t_k, which is received from the robot; $\mathcal{S}(i)$ is the set of neighbors of the i-th UWB anchor; $d_{e_{ij}}$ is the echo measurement between the i-th and the j-th UWB anchors; ${}^G\mathbf{p}_{a_i}$ and ${}^G\mathbf{p}_{a_j}$ are the global positions of the i-th and the j-th UWB anchors, which are unknown and need to be optimized; σ_r and σ_e are noise densities of ranging and echo measurements.

Following [19], we solve (10) with the distributed alternating direction method of multipliers (D-ADMM). In the D-ADMM algorithm, we associate a dual variable $\boldsymbol{\lambda}_{ij}$ with each constraint $\mathbf{x}_i^c = \mathbf{x}_j^c$ ($i, j = 1, \cdots, n$). Each *Opt. Node* in Fig. 2 keeps a local decision estimate $\mathbf{x}_i^c$ and a vector of dual variables $\lambda_{\mathbf{ji}}$ with $j < i$. Each *Opt. Node* in Fig. 2 updates the local state and the dual variables alternatively:

$$\mathbf{x}_{i,k+1}^c = \arg\min_{\mathbf{x}_i^c} f_i(\mathbf{x}_i^c) + \frac{\beta}{2}\sum_{j\in\mathcal{P}(i)} \|\mathbf{x}_{j,k+1}^c - \mathbf{x}_i^c - \frac{1}{\beta}\boldsymbol{\lambda}_{ji}^k\|^2 + \frac{\beta}{2}\sum_{j\in\mathcal{S}(i)} \|\mathbf{x}_i^c - \mathbf{x}_{j,k}^c - \frac{1}{\beta}\boldsymbol{\lambda}_{ij}^k\|^2 \tag{12}$$

$$\boldsymbol{\lambda}_{ji}^{k+1} = \boldsymbol{\lambda}_{ji}^k - \beta\,(\mathbf{x}_{j,k+1}^c - \mathbf{x}_{i,k+1}^c), \quad \forall j \in \mathcal{P}(i) \tag{13}$$

where $f_i(\mathbf{x}_i^c)$ is formulated as (11); $\beta > 0$ is a scalar that penalties the consensus constraints; $\boldsymbol{\lambda}_{ji}^k$ ($j \in \mathcal{P}(i)$) are dual variables preserved by the i-th *Opt. Node*; $\boldsymbol{\lambda}_{ij}^k$ ($j \in \mathcal{S}(i)$) are dual variables received from the successors of the i-th *Opt. Node*; $\mathcal{P}(i)$ and $\mathcal{S}(i)$ are the predecessor set and the successor set of the i-th *Opt. Node*.

The *Opt. Nodes* are ordered by the UWB anchors' IDs. At each iteration, the *Opt. Nodes* update with two steps. First, they solve (12) with the Levenberg-Marquard (LM) to update the local decision estimates and further update the dual variables with (13). Second, they send the new decision estimates $\mathbf{x}_{i,k+1}^c$ to their successors and the new dual variables $\boldsymbol{\lambda}_{ji}^{k+1}$ to their predecessors for the next iteration.

IV. Consistent VIRO Filter

In this section, we describe in detail the filter-based First-Estimate Jacobian Visual-Inertial-Ranging Odometry (FEJ-VIRO) running on the robot, which fuses measurements from UWB, cameras, and IMU in an MSCKF framework. We only present the different parts with [5] and refer the readers to our previous work [5] for more details.

A. State Vector

The state vector at time t_k consists of the current IMU states, the position of SLAM features, the position of UWB

anchors, and a short window, which contains cloned IMU poses corresponding to the past images:

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{x}_{I,k}^\top & \mathbf{x}_a^\top & {}^G\mathbf{p}_f^\top & \mathbf{x}_{cl,k}^\top \end{bmatrix}^\top \tag{14}$$

$$\mathbf{x}_{I,k} = \begin{bmatrix} {}^I_G\bar{\mathbf{q}}_k^\top & \mathbf{b}_{g,k}^\top & {}^G\mathbf{v}_{I,k}^\top & \mathbf{b}_{a,k}^\top & {}^G\mathbf{p}_{I,k}^\top \end{bmatrix}^\top \tag{15}$$

$$\mathbf{x}_a = \begin{bmatrix} {}^G\mathbf{p}_{a_1} & \cdots & {}^G\mathbf{p}_{a_n} \end{bmatrix}^\top \tag{16}$$

$$\mathbf{x}_{cl,k} = \begin{bmatrix} {}^I_G\bar{\mathbf{q}}_k^\top & {}^G\mathbf{p}_{I,k}^\top & \cdots & {}^I_G\bar{\mathbf{q}}_{k-m+1}^\top & {}^G\mathbf{p}_{I,k-m+1}^\top \end{bmatrix}^\top \tag{17}$$

where ${}^G\mathbf{p}_{a_i,k}$ $(i = 1, \cdots, n)$ are the positions of the UWB anchors expressed in $\mathcal{F}_G$; other variables have the same meaning with that of [5].

B. Ranging Update

Following [28], the UWB ranging measurement from the UWB anchor, d_r, can be modeled as:

$$\begin{aligned} d_r &= h(\mathbf{x}_I, {}^G\mathbf{p}_a) + n_r \\ &= \|{}^G\mathbf{p}_I + {}^I_G\mathbf{R}^\top {}^I\mathbf{p}_r - {}^G\mathbf{p}_a\| + d_{bias} + n_r \end{aligned} \tag{18}$$

where $\{{}^I_G\mathbf{R}, {}^G\mathbf{p}_I\}$ is the IMU pose; ${}^I\mathbf{p}_r$ is the position of the UWB node expressed in $\mathcal{F}_I$; ${}^G\mathbf{p}_a$ is the position of the UWB anchor expressed in $\mathcal{F}_G$; n_r is the zero-mean Gaussian noise; and d_{bias} is the bias of distance measurements, which can be calibrated offline [29].

We define the unbiased distance estimate as $\hat{d} = \|{}^G\hat{\mathbf{p}}_I + {}^I_G\hat{\mathbf{R}}^\top {}^I\mathbf{p}_r - {}^G\hat{\mathbf{p}}_a\|$ for concise representation. Then we perturb d_r to derive the measurement Jacobians:

$$\frac{\partial \tilde{d}_r}{\partial\, {}^G\tilde{\mathbf{p}}_I} = \frac{1}{\hat{d}} \cdot ({}^G\hat{\mathbf{p}}_r - {}^G\hat{\mathbf{p}}_a)^\top \tag{19}$$

$$\frac{\partial \tilde{d}_r}{\partial\, {}^I_G\tilde{\boldsymbol{\theta}}} = \frac{1}{\hat{d}} \cdot ({}^G\hat{\mathbf{p}}_a - {}^G\hat{\mathbf{p}}_u)^\top \cdot {}^I_G\hat{\mathbf{R}}^\top \lfloor {}^I\mathbf{p}_r \times \rfloor \tag{20}$$

$$\frac{\partial \tilde{d}_r}{\partial\, {}^G\tilde{\mathbf{p}}_a} = \frac{1}{\hat{d}} \cdot ({}^G\hat{\mathbf{p}}_a - {}^G\hat{\mathbf{p}}_r)^\top \tag{21}$$

where ${}^G\hat{\mathbf{p}}_r = {}^G\hat{\mathbf{p}}_I + {}^I_G\hat{\mathbf{R}}^\top {}^I\mathbf{p}_r$ is the position of UWB ranging node represented in the global frame.

C. Covariance Recovery with CI

Although the covariance recovery with CI runs on the newly-initialized UWB anchors, we describe it in this module because its goal is to maintain the consistency of the VIRO filter when augmenting the initialized anchor position in the robot state. The consistency is ensured by recovering the upper bound of the anchor position's covariance $\mathbf{P}_{aa}$.

We assume that M robot poses are used to initialize the i-th UWB anchor's position, and we denote these poses with:

$$\mathbf{x}_{rp} = \begin{bmatrix} {}^I_G\bar{\mathbf{q}}_1^\top & {}^G\mathbf{p}_{I,1}^\top & \cdots & {}^I_G\bar{\mathbf{q}}_M^\top & {}^G\mathbf{p}_{I,M}^\top \end{bmatrix}^\top \tag{22}$$

The robot also sends the marginal covariance of these poses $\Omega = \{\mathbf{P}_1 \ \cdots \mathbf{P}_M\}$, where $\mathbf{P}_i$ $(i = 1, \cdots, M)$ is 6×6 covariance of ${}^I_G\bar{\mathbf{q}}_i^\top$, ${}^G\mathbf{p}_{I,i}^\top$.

The corresponding M ranging measurements are defined to be $\mathbf{d}_{r_i} = [d_{r_i,1} \ \cdots \ d_{r_i,M}]^\top$, where $d_{r_i,k} = h_k(\mathbf{x}_{rp}, {}^G\mathbf{p}_{a_i}) + n_{i,k}$ $(k = 1, \cdots, M)$. Following [30] [5], we can have the following form by perturbing the observation module of ranging measurements and Givens-rotation:

$$\begin{bmatrix} \mathbf{r}_{i1} \\ \mathbf{r}_{i2} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{x1} \\ \mathbf{H}_{x2} \end{bmatrix} \tilde{\mathbf{x}}_{rp} + \begin{bmatrix} \mathbf{H}_{a1} \\ \mathbf{0} \end{bmatrix} {}^G\tilde{\mathbf{p}}_{a_i} + \begin{bmatrix} \mathbf{n}_{i1} \\ \mathbf{n}_{i2} \end{bmatrix} \tag{23}$$

Then we can evaluate the covariance matrix $\mathbf{P}_{aa}$ of the estimated anchor position ${}^G\hat{\mathbf{p}}_{a_i}$ by:

$$\mathbf{P}_{aa} = \mathbf{H}_{a1}^{-1}(\mathbf{H}_{x1}\mathbf{P}_{xx}\mathbf{H}_{x1}^\top + \sigma_r^2\mathbf{I})\mathbf{H}_{a1}^{-\top} \tag{24}$$

where σ_r is the noise density of ranging measurements; $\mathbf{P}_{xx}$ is the full covariance matrix of robot poses $\mathbf{x}_{rp}$. Note that the $\Omega = \{\mathbf{P}_1 \ \cdots \mathbf{P}_M\}$ are the 6×6 diagonal sub-matrices of $\mathbf{P}_{xx}$. From (24), we need the upper bound of $\mathbf{P}_{xx}$ to compute the upper bound of $\mathbf{P}_{aa}$. From the theory of covariance intersection [25], the upper bound of $\mathbf{P}_{xx}$ is given by:

$$\hat{\mathbf{P}}_{xx} = \begin{bmatrix} \frac{1}{\omega_1}\mathbf{P}_1 & & \\ & \ddots & \\ & & \frac{1}{\omega_M}\mathbf{P}_M \end{bmatrix} > \mathbf{P}_{xx} \tag{25}$$

where $\omega_1, \cdots, \omega_M$ are CI coefficients satisfying $\sum_{i=1}^M \omega_i = 1$ with $\omega_i > 0$. A tight upper bound of $\hat{\mathbf{P}}_{xx}$ can be obtained by adjusting the CI weights [31] [32].

By subscribing (25) in (24), we can have an upper bound of the anchor covariance $\hat{\mathbf{P}}_{aa}$. If $\{{}^G\hat{\mathbf{p}}_{a_i}, \hat{\mathbf{P}}_{aa}\}$ has not been augmented in the robot state, we just augment it to the state and the covariance. If the state already contains its estimated position ${}^G\mathbf{p}_{a_i,s}$ with covariance $\mathbf{P}_{aa,s}$, we fuse the existing value and the newly estimated value with CI:

$$\bar{\mathbf{P}}_{aa}^{-1} = \omega\hat{\mathbf{P}}_{aa}^{-1} + (1-\omega)\mathbf{P}_{aa,s}^{-1} \tag{26}$$

$$\bar{\mathbf{P}}_{aa}^{-1}{}^G\bar{\mathbf{p}}_{a_i} = \omega\hat{\mathbf{P}}_{aa}^{-1}{}^G\hat{\mathbf{p}}_{a_i} + (1-\omega)\mathbf{P}_{aa,s}^{-1}\,{}^G\mathbf{p}_{a_i,s} \tag{27}$$

where $\omega \in [0, 1]$ is the CI coefficient; $\{{}^G\bar{\mathbf{p}}_{a_i}, \bar{\mathbf{P}}_{aa}\}$ is the fused position and covariance.

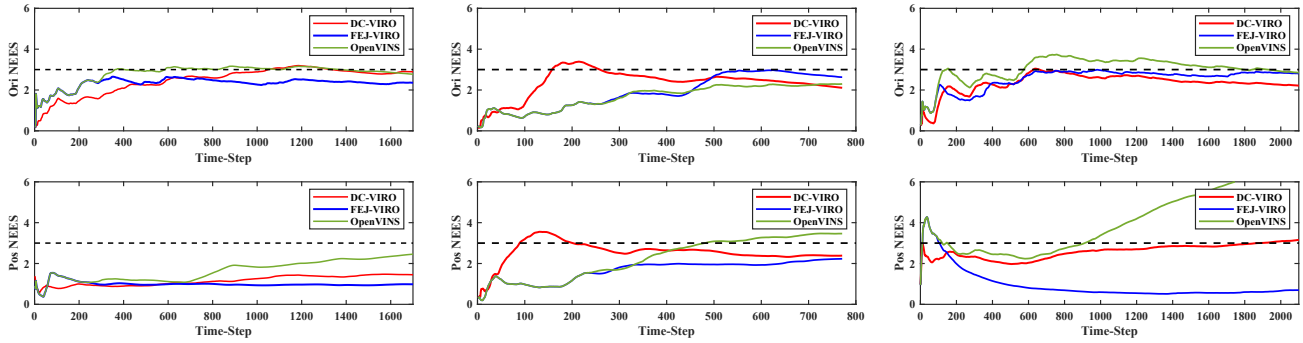
V. Experiments

We implement the *Consistent VIRO Filter* (see Fig. 1) based on *OpenVINS* [1], which is the state-of-the-art filter based visual-inertial estimator. The distributed UWB initialization method is implemented based on the Robot Operating System (ROS).

A. Simulation Experiments

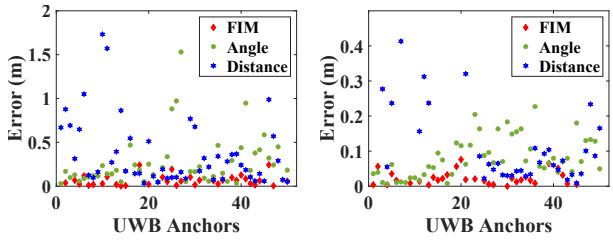
We expanded the simulation utility provided by *OpenVINS* to additionally simulate ranging measurements from $8 \sim 12$ UWB anchors with a noise density of *0.15m* and bias of *-0.75m*. The distance between two UWB sensors can be measured when below *20m*. We compare the NEES (see p31 of [33]) of the proposed *DC-VIRO* with *OpenVINS* and our previous *FEJ-VIRO* [5] to evaluate the consistency of the proposed system. The NEES results are listed in Fig.4. The dotted line in Fig.4 represents the 3 degree-of-freedoms of the position and orientation. The results show that our system is consistent.

For the initialization accuracy of UWB anchors, we compare two kinds of robot trajectories: *planar motion* and *3d random motion*. In our previous work [5], we used a



(a) (b) (c)

Fig. 4: Normalized estimation error squared (NEES) results of *DC-VIRO*, *FEJ-VIRO*, *OpenVINS* tested on three different simulation datasets.



(a) Planar Error (m) (b) 3D Random Error (m)

Fig. 5: Initialization errors of *50* randomly generated UWB anchor when the robot performing *planar motion* and *3D random motion*. Three different criteria are compared.

TABLE I: Average ATE (m) of *six* runs in VIRAL Dataset

	OpenVINS	VINS-Fusion (RT / LC)	DC-VIRO (Distributed)	FEJ-VIRO (Centralized)
eee_01	0.647	0.564 / *0.364*	0.524	**0.479**
eee_02	0.604	0.582 / *0.252*	0.382	**0.340**
eee_03	0.417	0.540 / 0.390	**0.331**	0.348
nya_01	0.741	0.538 / *0.235*	**0.412**	0.505
nya_02	0.478	0.375 / 0.331	0.217	**0.202**
nya_03	0.692	0.748 / 0.370	**0.263**	0.290

naive criterion, which considers the total moving distance of the robot and optimizes the UWB anchors' positions when the total distance is over a threshold. In [22], the authors proposed increasing the spanning angles between the consecutive measurements to obtain accurate UWB localization. So we implemented three different criteria for UWB anchors' initialization: *distance-based criterion* with threshold *15m*, *angle-based criterion* with threshold *0.3 rad*, and our *FIM-based criterion* with threshold *8000*. Note that the determinant of FIM has no unit. Each *Opt. Node* only preserves *30* ranging measurements and the corresponding robot poses. The *Opt. Node* performs a uniform down-sample to keep a bounded number of measurements. The initialization errors of UWB anchor initialization are presented in Fig. 5 for planar robot motion (a) and 3D random motion (b). Fig. 5 shows that the errors of our *FIM-based criterion* are significantly smaller than that of the other two criteria when the robot moves on a planar, while *FIM-based criterion* and *angle-based criterion* performs better than the *distance-based criterion* for the 3D random motion condition.

B. Real-world Experiments

We further evaluate our method based on a real-world dataset, VIRAL [28], which provides measurements from extensive sensors, such as IMU, stereo cameras, and *three* UWB anchor nodes. For the VIRO filter, We use 11 clones and 200 features for real-time estimation.

We test the proposed *DC-VIRO* on six trajectories available in the VIRAL dataset. And we compare the results with the centralized *FEJ-VIRO* [5], the optimization-based *VINS-Fusion* and the original *OpenVINS*. Experiment results are listed in Tab.I. The column of *VINS-Fusion (RT)* represents the *real-time pose estimates*, while the column of *VINS-Fusion (LC)* represents the *optimized path* after loop-closure.

From Tab. I, our *DC-VIRO* is accurate and is nearly as accurate as the *optimized loop-closure path* of *VINS-Fusion* and the centralized method *FEJ-VIRO*. Note that the *DC-VIRO* achieves higher accuracy than the *FEJ-VIRO* in *eee_03*, *nya_01*, and *nya_03*. This is because these trajectories contain more straight lines or planar motion, and the *FIM-based criterion* helps to better initialize the UWB anchors' positions, leading to higher estimation accuracy. While the other three trajectories provide more 3D random motions, the *FEJ-VIRO* can initialize UWB anchors' positions well even with the naive *distance-based criterion*. On the other hand, the long-window structure of the centralized *FEJ-VIRO* allows better covariance recovery, thus leading to higher accuracy when there are only *three* UWB anchors.

VI. CONCLUSIONS

In this paper, we propose a distributed system that fuses measurements from cameras, IMU, and UWB consistently to estimate robot pose and the UWB anchor positions. The robot platform and all UWB anchors equipped with an onboard computer are computing nodes in a cluster, and only state values and marginal covariance are transmitted between them. The robot runs a lightweight and consistent VIRO filter to estimate the robot poses. While the UWB network estimates the anchor positions with D-ADMM once the FIM-based criterion is met. The estimated anchor positions are sent to the robot and augmented into the VIRO filter consistently with the CI technique.

For future work, we would try to extend the DC-VIRO to a multi-robot system.

REFERENCES

[1] P. Geneva, K. Eckenhoff, W. Lee, Y. Yang, and G. Huang, "OpenVINS: A research platform for visual-inertial estimation," in *Proc. of the IEEE International Conference on Robotics and Automation*, Paris, France, 2020. [Online]. Available: https://github.com/rpng/open_vins

[2] A. Mourikis and S. Roumeliotis, "A multi-state constraint kalman filter for vision-aided inertial navigation," in *2007 IEEE International Conference on Robotics and Automation, ICRA'07*, ser. Proceedings - IEEE International Conference on Robotics and Automation, 2007, pp. 3565–3572, copyright: Copyright 2011 Elsevier B.V., All rights reserved.; 2007 IEEE International Conference on Robotics and Automation, ICRA'07 ; Conference date: 10-04-2007 Through 14-04-2007.

[3] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based visual–inertial odometry using nonlinear optimization," *The International Journal of Robotics Research*, vol. 34, no. 3, pp. 314–334, 2015.

[4] G. Huang, "Visual-inertial navigation: A concise review," in *2019 international conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 9572–9582.

[5] S. Jia, Y. Jiao, Z. Zhang, R. Xiong, and Y. Wang, "Fej-viro: A consistent first-estimate jacobian visual-inertial-ranging odometry," *arXiv preprint arXiv:2207.08214*, 2022.

[6] T. H. Nguyen, T.-M. Nguyen, and L. Xie, "Tightly-coupled single-anchor ultra-wideband-aided monocular visual odometry system," in *2020 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2020, pp. 665–671.

[7] B. Yang, J. Li, and H. Zhang, "Uvip: Robust uwb aided visual-inertial positioning system for complex indoor environments," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 5454–5460.

[8] J. Tiemann, A. Ramsey, and C. Wietfeld, "Enhanced uav indoor navigation through slam-augmented uwb localization," in *2018 IEEE international conference on communications workshops (ICC workshops)*. IEEE, 2018, pp. 1–6.

[9] B. Yang, J. Li, and H. Zhang, "Resilient indoor localization system based on uwb and visual-inertial sensors for complex environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–14, 2021.

[10] A. Saipulla, B. Liu, and J. Wang, "Barrier coverage with airdropped wireless sensors," in *MILCOM 2008-2008 IEEE Military Communications Conference*. IEEE, 2008, pp. 1–7.

[11] M. R. Cacan, E. Scheuermann, M. Ward, M. Costello, and N. Slegers, "Autonomous airdrop systems employing ground wind measurements for improved landing accuracy," *IEEE/ASME Transactions on Mechatronics*, vol. 20, no. 6, pp. 3060–3070, 2015.

[12] P. Lutz, M. J. Schuster, and F. Steidle, "Visual-inertial slam aided estimation of anchor poses and sensor error model parameters of uwb radio modules," in *2019 19th International Conference on Advanced Robotics (ICAR)*. IEEE, 2019, pp. 739–746.

[13] T.-M. Nguyen, S. Yuan, M. Cao, T. H. Nguyen, and L. Xie, "Viral slam: Tightly coupled camera-imu-uwb-lidar slam," *arXiv preprint arXiv:2105.03296*, 2021.

[14] Y. Xiong, N. Wu, Y. Shen, and M. Z. Win, "Cooperative localization in massive networks," *IEEE Transactions on Information Theory*, vol. 68, no. 2, pp. 1237–1258, 2021.

[15] S. Srirangarajan, A. H. Tewfik, and Z.-Q. Luo, "Distributed sensor network localization using socp relaxation," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4886–4895, 2008.

[16] A. Simonetto and G. Leus, "Distributed maximum likelihood sensor network localization," *IEEE Transactions on Signal Processing*, vol. 62, no. 6, pp. 1424–1437, 2014.

[17] R. M. Buehrer, H. Wymeersch, and R. M. Vaghefi, "Collaborative sensor network localization: Algorithms and practical issues," *Proceedings of the IEEE*, vol. 106, no. 6, pp. 1089–1114, 2018.

[18] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[19] E. Wei and A. Ozdaglar, "Distributed alternating direction method of multipliers," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012, pp. 5445–5450.

[20] T.-H. Chang, M. Hong, W.-C. Liao, and X. Wang, "Asynchronous distributed admm for large-scale optimization—part i: Algorithm and convergence analysis," *IEEE Transactions on Signal Processing*, vol. 64, no. 12, pp. 3118–3130, 2016.

[21] Z. Wang and G. Dissanayake, "Observability analysis of slam using fisher information matrix," in *2008 10th International Conference on Control, Automation, Robotics and Vision*. IEEE, 2008, pp. 1242–1247.

[22] T. H. Nguyen and L. Xie, "Estimating odometry scale and uwb anchor location based on semidefinite programming optimization," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7359–7366, 2022.

[23] ——, "Relative transformation estimation based on fusion of odometry and uwb ranging data," *arXiv preprint arXiv:2202.00279*, 2022.

[24] S. Julier and J. K. Uhlmann, "General decentralized data fusion with covariance intersection," in *Handbook of multisensor data fusion*. CRC Press, 2017, pp. 339–364.

[25] L. Chen, P. O. Arambel, and R. K. Mehra, "Estimation under unknown correlation: Covariance intersection revisited," *IEEE Transactions on Automatic Control*, vol. 47, no. 11, pp. 1879–1882, 2002.

[26] T.-K. Chang, K. Chen, and A. Mehta, "Resilient and consistent multirobot cooperative localization with covariance intersection," *IEEE Transactions on Robotics*, vol. 38, no. 1, pp. 197–208, 2021.

[27] 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)*, 2016, pp. 4289–4296.

[28] T.-M. Nguyen, S. Yuan, M. Cao, Y. Lyu, T. H. Nguyen, and L. Xie, "Ntu viral: A visual-inertial-ranging-lidar dataset, from an aerial vehicle viewpoint," *Internationl Journal of Robotics Research (accepted, to appear)*, 2021.

[29] B. van der Heijden, A. Ledergerber, R. Gill, and R. D'Andrea, "Iterative bias estimation for an ultra-wideband localization system," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 1391–1396, 2020.

[30] M. Li, *Visual-inertial odometry on resource-constrained systems*. University of California, Riverside, 2014.

[31] W. Niehsen, "Information fusion based on fast covariance intersection filtering," in *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002.(IEEE Cat. No. 02EX5997)*, vol. 2. IEEE, 2002, pp. 901–904.

[32] D. Franken and A. Hupper, "Improved fast covariance intersection for distributed data fusion," in *2005 7th International Conference on Information Fusion*, vol. 1. IEEE, 2005, pp. 7–pp.

[33] K. J. K. Mark L. Psiaki. Model-based estimation notes. [Online]. Available: http://www.mit.edu/ kircher/mbe.pdf