2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

Robust Indoor Localization with Ranging-IMU Fusion

Fan Jiang[1,2,*], David Caruso[1], Ashutosh Dhekne[2], Qi Qu[1], Jakob Julian Engel[1], Jing Dong[1]

***Abstract*— Indoor wireless ranging localization is a promising approach for low-power and high-accuracy localization of wearable devices. A primary challenge in this domain stems from non-line of sight propagation of radio waves. This study tackles a fundamental issue in wireless ranging: the unpredictability of real-time multipath determination, especially in challenging conditions such as when there is no direct line of sight. We achieve this by fusing range measurements with inertial measurements obtained from a low cost Inertial Measurement Unit (IMU). For this purpose, we introduce a novel asymmetric noise model crafted specifically for non-Gaussian multipath disturbances. Additionally, we present a novel Levenberg-Marquardt (LM)-family trust-region adaptation of the iSAM2 fusion algorithm, which is optimized for robust performance for our ranging-IMU fusion problem. We evaluate our solution in a densely occupied real office environment. Our proposed solution can achieve temporally consistent localization with an average absolute accuracy of ∼0.3m in real-world settings. Furthermore, our results indicate that we can achieve comparable accuracy even with infrequent range measurements down to 1Hz.**

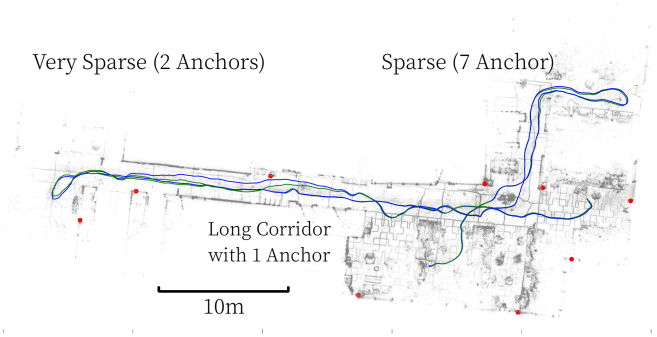

Fig. 1. Localizing in a cluttered, real office environment with sparsely populated anchors using proposed approach, with only UWB ranging and IMU sensor data. The average/max trajectory error are 0.3/1.0 meters. Green: ground truth trajectory. Blue: result trajectory. Red: UWB anchors. Note while the solution quality degrades in the corridor, due to very few line-of-sight ranging measurements in the corridor, it is able to recover once enough measurements are available.

I. Introduction

Indoor localization is a core requirement for virtual reality/augmented reality (VR/AR) devices and robots. Traditionally, accurate indoor 6 Degrees-of-Freedom (DoF) localization is performed using Visual-Inertial Odometry (VIO) [1], and visual Simultaneous Localization and Mapping (SLAM) [2], [3], [4], in which camera images are the main source of information. However, exposing raw images of the surrounding environment to localization algorithms has privacy implications, and capturing and processing a large amount of raw image pixels lead to significant energy consumption and heat dissipation, which is not ideal for small form factor wearable devices.

In recent years, wireless ranging based indoor localization has attracted significant research attention, due to its good accuracy and low power consumption. For example, the average ranging power of an Ultra-Wide Band (UWB) system could be as low as 0.14mW [5], assuming 1Hz ranging frequency (2.5ms per ranging, 55mW average power during ranging), which fits well within the power envelope of wearable devices such as smart glasses and watches.

While promising, wireless ranging is not free of issues. In particular, non-line of sight (NLOS) propagation of wireless signals, where radios waves undergo reflections from surrounding objects, can deteriorate ranging accuracy. In indoor environments cluttered with objects, NLOS measurements can become dominant, making conventional Gaussian noise models ill-suited. Even more complex scenarios arise when a user lacks line-of-sight to any localization transceivers, a situation frequently encountered in multi-room setups.

Various hardware and signal processing methodologies have been proposed by the wireless community to mitigate multipath in ranging measurements. Notable strategies include using Channel Impulse Response (CIR) [6], multiple antennas (Multi-Input-Multi-Output, MIMO) [7], beamforming [8] and angle-of-arrival [7]. On the other hand, the state estimation community tries to address the same problem by fusing together wireless ranging measurements with other sensors, to reduce ambiguity in multi-path determination during measurements [9], [10], [11].

Depending on the current estimate's uncertainty, sufficient data may not be available for identifying NLOS measurement causally at all time, so the system needs to keep a long horizon of previous nonlinear information and correct past states as necessary. Such a requirement aligns seamlessly with incremental smoothing techniques like [12].

In this work, we present a novel solution to the ranging-IMU fusion problem. Our main contributions are:

1) Designed, built and evaluated a system that can robustly localize a custom device composed of an IMU and UWB receiver in real-time, even in challenging environments.
2) A novel asymmetric m-Estimator for wireless Time-of-Flight (ToF) ranging measurement, to correctly model non-Gaussian multi-path effects;
3) A practical way to improve the numerical stability of

*This work was mostly done during the internship at Meta.
[1]Meta Reality Labs Research. {dcaruso, qqu, jakob.engel, jingdong}@meta.com
[2]Georgia Institute of Technology. {fan.jiang, dhekne}@gatech.edu

iSAM2, with presence of strong local non-linearity, without using complex trust region update strategies like Dogleg;

We demonstrate that even with single-sided two-way ranging (SS-TWR) and sparsely placed UWB anchors we can achieve consistent indoor localization with range measurements with a mean absolute accuracy of $\sim$0.3m, and maintain reasonable location accuracy even with 1Hz measurements (see Figure 1).

The rest of the paper is organized as follows: we discuss related work in Sec. II followed by the model used for the sensors in Sec. III. In particular, it includes the model used for wireless ranging measurements: the main contribution of the paper. The improvement of the robustness of iSAM2, is described in Sec. IV, which forms the second contribution of this work. We finally discuss the implementation details of our prototype and our evaluation results in Sec. VI.

II. Related Work

A. Related Work in Wireless Signal Processing

Many methods have been proposed to reduce the effect to multipath and improve the probability of detecting the "first path" (the shortest propagation path) in complex indoor environments, for example with neural networks [13] and with signal processing techniques [14], sometimes even with multiple antennae (Multi-Input-Multi-Output, MIMO) [7]. A notable strategy utilizes CIR of received signals to discern the direct-path measurements [6], [15], where each "peak" in the energy part of the CIR corresponds to a different propagation path of the original transmitted signal in the space between the transmitter and the receiver. The idea is to use the CIR to identify the first path in environments with complex multipath effects. However, this approach is not robust in real world applications since path resolution below the limitations imposed by the wireless signal's bandwidth is challenging. The CIR often provides ambiguous information without knowing the detailed environment structure. For example, when the direct path is blocked by walls, or has very low signal strength, any path identified by these approaches will lead to wrong range measurements. Efforts using signal processing to identify such edge cases can lead to excessive computation and diminishing returns [16]. While techniques like beamforming [8] and angle-of-arrival [7] provide supplementary insights, they often need additional hardware, and yet the fundamental challenges remain.

B. Related Work in State Estimation

Existing work try to solve multipath problems with sensor fusion, where the ranging measurements are fused with a interoceptive motion sensor, usually an Inertial Measurement Unit (IMU). Some of the first works in this space use Extended Kalman Filter (EKF) and prediction error based outlier rejection ([9], [10], [11]). After the proposal of the pre-integration technique [17], researchers have used smoothing and mapping techniques, for example factor graph optimization with iSAM2 [18]. However, filtering approaches using IMU fusion without utilizing past data can only identify sparse outliers in range measurements. Such techniques cannot recover when valid ranging measurements are absent for significant time, since prediction is likely to drift away, worsening the problem for future valid measurements.

One possible way of modeling the effect of non-line of sight measurement is to capture the inlier/outlier ambiguity explicitly by using binary discrete variables in a *hybrid* factor graph. When incremental inference is required, explicit incremental solvers such as iMHS [19], NF-iSAM [20], or MH-iSAM2 [21] can be used. However, such hybrid modeling require mixed-integer program (MIP) solvers or Expectation-Maximization (EM) style solvers [22], whose cost could be prohibitive in real-time applications.

Another common way to model outliers is *m*-Estimator. m-Estimators have already been shown to map directly to E-M methods [23] in the continuous formulation. The *m*-Estimator for Cauchy distributions can be attributed to Barnett's 1966 work [24]. The frequently referenced article on m-Estimators [25], makes an assertion—without specific citation—that the Cauchy m-Estimator often produces incorrect results without a means of verifying their accuracy. While this assertion holds merit, it somewhat oversimplifies Barnett's original proposition. Authors of [24] specifically postulated that any local approach resembling Newtonian methods has the possibility of being unsuccessful, however such instances might be infrequent in practice. Moreover, he emphasized that evading local minima is unfeasible without a comprehensive exploration of the likelihood function—a statement universally applicable to all gradient-driven local methodologies. Our proposed approach is based on m-Estimators, and we will demonstrate with our modified half-Cauchy m-Estimator, our approach delivers promising results.

III. Ranging-IMU Measurement Model

A. IMU Model

We use a typical IMU model. Assuming zero noise and known initial condition and a flat earth approximation, the simplified strapdown mechanization equations are given by:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{R}_t &= \boldsymbol{R}_t[\boldsymbol{\omega}_t]_\times \\ \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{V}_t &= \boldsymbol{g} + \boldsymbol{R}_t\boldsymbol{a}_t \\ \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}_t &= \boldsymbol{V}_t \end{aligned} \tag{1}$$

where $\boldsymbol{R}_t$ is the rotation matrix, $\boldsymbol{V}_t$ the body velocity in earth frame, $\boldsymbol{X}_t$ the translation vector, $\boldsymbol{g}$ the gravity vector in the earth frame, $\boldsymbol{\omega}_t$ is current angular velocity in the IMU frame as measured by the gyroscope, and $\boldsymbol{a}_t$ the current total body acceleration in the IMU frame as measured by the accelerometer. After applying a factory calibration to the IMU signal, we assume the rectified signal to be polluted by a Gaussian noise and a slowly varying biases, 3 components for the accelerometer bias $\boldsymbol{b}_t^a$ and gyroscope axis $\boldsymbol{b}_t^g$. To model the later we chose a stochastic 1st-order Gauss-Markov random processes described here [26]

(Section 5.2.4):

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{b}_t = -\frac{1}{\tau_{\text{bias}}}\boldsymbol{b}_t + w_{\text{bias}}(t) \tag{2}$$

Where $\boldsymbol{b}_t = [\boldsymbol{b}_t^a, \boldsymbol{b}_t^q]^{\mathrm{T}}$ denotes both bias components of the IMU. τ_{bias} are a correlation time constant and w_{bias} random variables following a centered Gaussian distribution with (diagonal) covariance $\Sigma_{w_{\text{bias}}}$. The parameters of the model were derived from a recorded Allan Variance method slightly inflated to increase robustness to unmodeled effects.

This IMU model is leveraged with two types of factor in our optimization based smoother. The propagation constraint is using the *preintegration* technique as described in [17], [27]. The latter allows to easily write the likelihood of the preintegration measurement as:

$$L_{\text{preint}} = \exp\left(-\frac{1}{2}\left\|r_{\text{IMU}}\left((\boldsymbol{RVX})_{t+1}, (\boldsymbol{RVX})_t, \boldsymbol{b}_t\right)\right\|^2_{\Sigma_{\text{IMU}}}\right) \tag{3}$$

Where the residual r_{IMU} and its covariance Σ_{IMU} are defined respectively by Eq. 37 and Eq. 35 in [27] substituting i by t and j by $t+1$. The biases constraints are written with a binary factor between consecutive bias estimates which represents the likelihood:

$$L_{\text{bias}} = \exp\left(-\frac{1}{2}\left\|\boldsymbol{b}_{t+1} - e^{\frac{\Delta t_{t;t+1}}{\tau}}\boldsymbol{b}_t\right\|^2_{\Sigma_{w_{\text{bias}}}}\right) \tag{4}$$

B. Ranging Measurement Model

The wireless ranging measurement factor's likelihood is modeled with a mixture of two probability distributions. One distribution is for the line of sight measurements, where the noise mainly comes from the inaccuracies of the First Path Estimation (FPE) algorithm. This error is mainly caused by inaccuracies in the radio hardware, as well as the limited energy and bandwidth of the transmitted radio signal. This noise appears in the range measurement as a Gaussian-like noise of a σ of about 0.1-0.2m [28]. The other distribution is for the NLOS measurements, where the measurements follow an **environment-dependent** unknown distribution, which could be as large as twice of the actual distance, or as small as indistinguishable from the Gaussian measurement noise.

$$L_{\text{range}}(r|\bar{r}, m) \sim \begin{cases} \mathcal{N}(\bar{r}; \sigma_r), & m = 0 \\ L_{\text{range}}(r|\bar{r}, m = 1), & m = 1 \end{cases} \tag{5}$$

where r is the measured range $r = \|\boldsymbol{X}_i - \boldsymbol{A}_j\|_2$ where $\boldsymbol{A}_j$ is the anchor j's 3D position, $\bar{r}$ is the true range, σ_r is the range variance, and $m = \{0, 1\}$ is the discrete variable that indicates if the measurement is a direct path measurement or not, with $m = 0$ indicates the range measurement is direct, while $m = 1$ indicates it is not, which we will dwell on next.

Existing work [29] factorizes the NLOS distribution into two distributions with a convolution of the Gaussian and an exponential distribution, which is the maximum entropy distribution supported on $[0, +\infty)$ with some statistical moment, with a sample mean and variance obtained by simulation of the ranging process in a simulator. However, in wireless ranging problems we cannot apply the same model as in [29], since the environment is unknown and hence we have little information about $p(r|\bar{r}, m = 1)$,

Instead, as the first contribution of this paper, we model the NLOS measurements with a half-Cauchy distribution, in line with the maximum entropy principle, with location $\theta = 0$ and scale γ, supported also on $[0, +\infty)$:

$$p(r' - \bar{r}|m = 1) \sim \frac{2}{\pi\gamma}\frac{1}{1 + ((r' - r)/\gamma)^2} \tag{6}$$

where r' is the measured range, $\bar{r}$ the true range. The unknown scale parameter γ can be either determined by parameter estimation techniques using real or simulated data, or simply derived using the Interquartile Range of the (Gaussian) ranging noise heuristically.

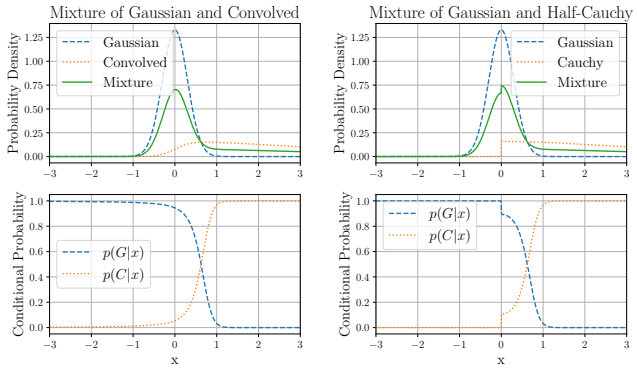

Fig. 2. Qualitative comparison of using (a) a combined marginal model and (b) a mathematically simpler half-Cauchy model for the multipath. Note the less complex model does not significantly change the decision point.

With this NLOS model, the observed range model for $m = 1$ is then (sum of independent variables is convolution)

$$L_{\text{range}}(r|\bar{r}, m = 1) = \int_{-\infty}^{+\infty} p(r|r')p(r' - \bar{r}|m = 1)\mathrm{d}r' \tag{7}$$

which is shown in Fig. 2a.

The physical intuition behind the use of a single-sided distribution for NLOS measurements, is that the true range is the shortest path in the space, and NLOS measurements are necessarily longer than the shortest path between anchor and receiver.

In Fig. 3 we show the evolution of the UWB ranging measurement error distribution across trajectories, using the ground-truth ranges computed through ground truth trajectories and anchor positions. The x axis is time elapsed, y axis is the range error in meters, and the color represents the distribution density. It is apparent that the distribution is single-sided, with occasional large outliers, and frequent smaller outliers.

It is important to note that while the marginal distribution for $r - \bar{r}$ may seem light-tailed, the actual distribution considering the hidden variable, the current surroundings of the user, is not. This can be observed on Fig. 3, where persistent NLOS measurements dominate in some time slices. This

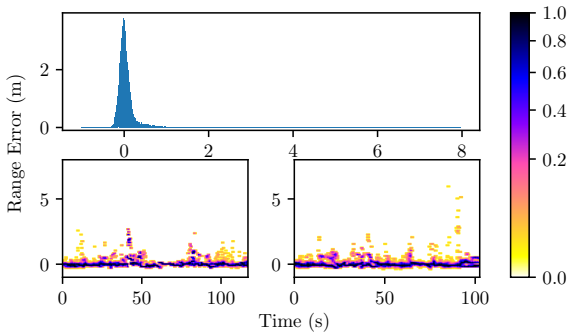

Fig. 3. Range measurement error distribution. Top: marginal over all trajectories seems light-tailed. Bottom: heat-map over time for run 1 and 2 of Table I. Note distribution becomes fat tailed in a lot of places, for example $t \sim 45s$ on the left, and $t \sim 80s$ on the right. Color indicates the value of the probability density function at time t.

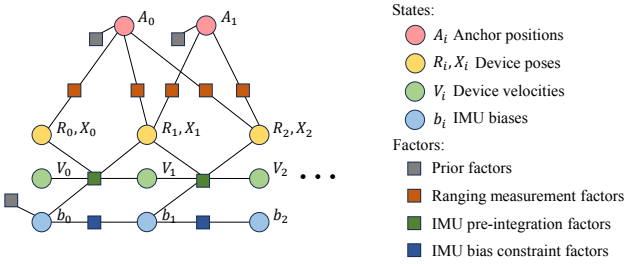

Fig. 4. An example factor graph of the ranging-IMU fusion.

explains why we need to use a half-Cauchy distribution without priors on m, instead of just using the measured range marginal density like in [29].

Other than considering explicit hybrid factor graph inferences to solve discrete variable m, we use an *implicit* approach to solve the range-IMU fusion problems, based on the m-Estimators to simplify the inference process. Since the decision boundary (where $m = 0$ or $m = 1$ is more likely) is the same, and the p.d.f. of both the marginal and the half-Cauchy distribution is very close after the decision boundary, we can use the half-Cauchy distribution directly as the mixture component (Fig. 2b).

The IMU and ranging model presented are used in a factor graph formulation. An example factor graph is shown in Fig. 4. The states include the devices pose + velocity + IMU bias at each time point when we receive a range measurement, and also positions of all rang-able anchors. The factors include the IMU pre-integration factors and ranging factors we just introduced, plus prior factors on first IMU bias with factory calibration, and all anchors with pre-mapped positions.

IV. Trust-Region Variant of iSAM2

A known issue of using iSAM2 framework in incremental inference is the occurrence of indeterminant linear systems [30]. This is primarily because iSAM2 internally uses a Gauss-Newton like update for solving the nonlinear least-squares problem, which is not robust to ill-conditioned problems.

While existing trust-region based methods like RISE [31] have been shown to perform well in regular SLAM problems, its convergence has not been validated with switchable methods like ours whose continuous error could change with the discrete decision variable. Same as previously reported [32], we observed that RISE does not work when the radius of the trust region changes. This leads to a performance similar to Gauss-Newton (vanilla iSAM2), which similarly fails to achieve optimal non-linear updates, resulting in worse performance in our application.

In contrast to the Dogleg-like algorithm proposed by [32] (which is concurrent to our work), as the second contribution of this paper, we propose **D-iSAM2**, a simpler trust-region method, which works well in our real-world experiments, and only requires minimal changes in iSAM2. This method shares the same core idea with the Levenberg-Marquadt (L-M) algorithm, whose linear update (in the simplest form) is calculated from

$$(J^{\mathrm{T}}J + \lambda I)\delta = J^{\mathrm{T}}z \tag{8}$$

where J is the Jacobian, λ the damping factor, δ the linear increment, z the linear error vector. This linear problem effectively is equivalent to solving a nonlinear factor graph with the following Jacobian structure

$$\begin{bmatrix} J^{\mathrm{T}} & \lambda^{1/2}I \end{bmatrix} \begin{bmatrix} J \\ \lambda^{1/2}I \end{bmatrix} \delta = \begin{bmatrix} J & \lambda^{1/2}I \end{bmatrix} \begin{bmatrix} z \\ 0 \end{bmatrix} \tag{9}$$

which is equivalent to the original graph with a *special* factor on each variable where the factor always has an error function of value 0, but a Jacobian of $\lambda^{1/2}I$. When the trust region is an ellipse, λI can be replaced by a diagonal of n lambdas, $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

When new information is added, the algorithm only needs to check whether the current λ is valid, by checking if the error decreases after the iSAM2 step with the special factor added. If not, increase λ, and if yes, decrease λ, then repeat the process. We refer the interested reader to [33] where the schedule for the λ modifications have been extensively covered.

Algorithm 1 D-iSAM2 Algorithm

Input: Bayes Tree $\mathcal{T}$, current estimate $\mathcal{X}$, new factors $\mathcal{F} = \{\phi_i\}$, new variable initial estimates $\{x_i\}$, current λ
  convergence indicator $c \leftarrow$ False
  while not converged c **do**
    $\mathcal{F}' \leftarrow$ AddSpecialFactor($\mathcal{F}$, λ)
    $\mathcal{T}', \mathcal{X}' \leftarrow$ iSAM2Update($\mathcal{T}, \mathcal{F}', \mathcal{X}, \{x_i\}$)
    if nonlinear error decreased **then**
      $c \leftarrow$ True
      Decrease λ with schedule ▷ until $\lambda_{\min}$
    else
      Increase λ with schedule
      if $\lambda = \lambda_{\max}$ throw error
    end if
  end while
Output: new Bayes Tree $\mathcal{T}'$, new estimate $\mathcal{X}'$

Note that the λ_{x_i} for an old variable x_i will not change after a successful D-iSAM2 step. This is intentional, since it

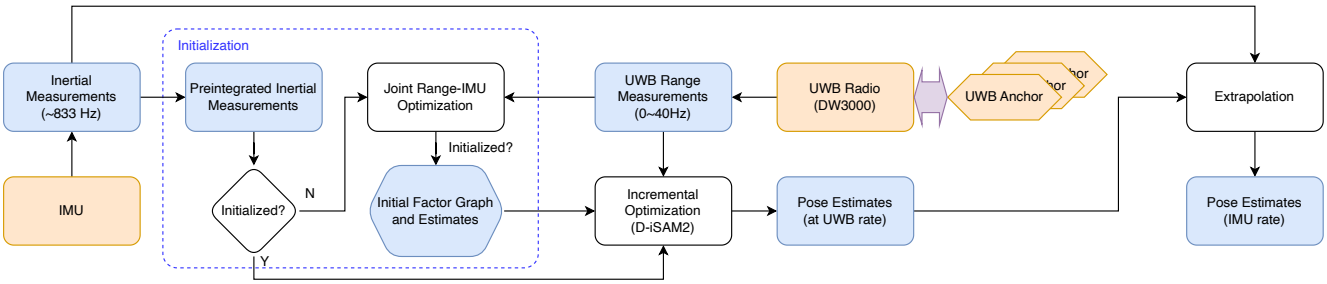

Fig. 5. System setup of our inertial-ranging fusion system. White blocks are algorithm, yellow blocks are hardware, and blue blocks are data.

is very costly to update all old Bayes tree nodes for a new set of $\{\lambda_{x_n}\}$, since all nodes will need to be re-factorized. However, since we operate incrementally, the last added variable is the most numerically challenging variable. Hence, only changing the trust region radius for the last variable is sufficient. The whole algorithm is described in Alg. 1.

Changing the λ for past variables, which we explicitly choose not to do, is a "global" process similar to full relinearization in iSAM2. The process impacts every node in the factor graph, hence could be detrimental to the overall performance if done in a naive way. However, this may be required in some problems more difficult than ours. Also, child nodes can be marginalized once they are too old for the current estimate, like in a fixed lag smoother. Since this is not the primary aim of this paper, we will leave how to further optimize this trust-region based robust incremental optimization method for a future work.

V. Implementation

The overall system diagram is shown in Figure 5. Input sensor data of the system include an IMU operated at high frequency, and a UWB transceiver that executes SS-TWR range measurements from fixed anchors in the space. During the initialization phase, a joint of factor graph of ranging data and preintegration of the IMU data is built for 10 seconds, and a batch optimization of the factor graph is performed to obtain initial values of the system states, to initialize an iSAM2 estimator. After initialization, the raw ranging measurements and preintegrations are directly fed into iSAM2 to get UWB-frequency estimation, and IMU-frequency estimation can be future obtained by extrapolation using IMU measurements.

We implemented the wireless hardware using the Decawave DWM3000 module with ESP32-S3 as the main microcontroller. For the ranging protocol, we use a simple SS-TWR ranging scheme, where the device sends 1 packet to the anchor being ranged against and receive 1 packet with timestamp. We use SS-TWR because it possesses similar noise characteristics with Wi-Fi Fine Time Measurement (FTM), which enables our method to also apply to Wi-Fi localization. The IMU stream is directly recorded off the left IMU of the Project Aria device [34], which is a factory calibrated BMI263 from Bosch operating at 800Hz. The IMU stream is time synchronized with the UWB data, through a hardware time synchronization link between the UWB module and Project Aria device.

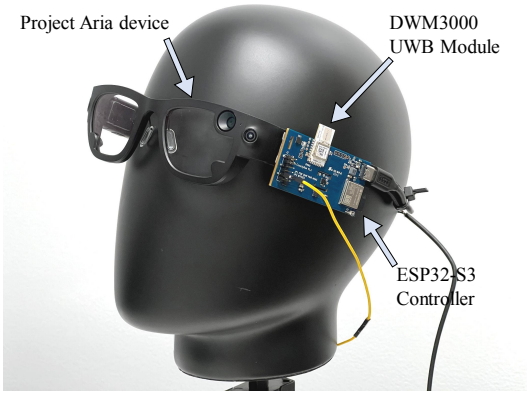

Fig. 6. Prototype hardware built, with a wireless module attached to a Project Aria device [34].

The fusion algorithm is implemented with C++ using the GTSAM [35] library using the preintegrated IMU factor of Sec. III-A. All evaluations are conducted on a Macbook Pro with an Apple M1 Pro chip.

VI. Evaluation

We evaluate the proposed system via experiments in a typical 30 by 50 meters office environment. Ground truth device trajectory is obtained with Project Aria Machine Perception Service [34], using Project Aria device's collected data as input. UWB anchors are co-located in the the ground truth trajectory frame of reference using 2D fiducials, with $< 1\,\mathrm{cm}$ accuracy. We evaluate the performance of our approach over 7 runs, each of which is a walk 1-3 minutes in duration and 50-130 meters long in travelled distance. The resulting trajectory is evaluated using the Evo library [36]. We use the mean Absolute Position Error (APE) as the metric. Note that this is different with SLAM/VIO, where APE is proportional to distance travelled. In our scenario all anchors are fixed.

Ranging frequency of each anchor is set at maximum 10Hz, and the receiver is configured to output measurements from up to 4 UWB anchors which have strongest signal strength, so the cumulative received ranging measurements are up to 40Hz. We will use 40Hz ranging frequency in Section VI-A, and explore lower frequency in VI-B.

TABLE I

3D ABSOLUTE POSITION ERROR (METERS) WITH DIFFERENT RANGING NOISE MODEL AND ESTIMATORS (AVERAGE/MAX), USING 40HZ RANGE MEASUREMENTS. THE BEST METHOD IS INDICATED IN HIGHLIGHTED WITH 5% TOLERANCE.

Noise model	Estimator	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7
Gaussian	D-iSAM2	0.68/4.44	0.90/2.28	0.59/1.98	0.43/0.98	1.08/10.43	**0.33/1.30**	0.66/6.06
Huber	D-iSAM2	0.54/4.02	0.35/1.83	0.52/**1.15**	**0.30/1.31**	0.45/5.05	0.70/2.57	0.41/1.31
Cauchy	D-iSAM2	0.36/1.27	0.78/1.62	0.62/1.74	**0.30**/1.40	0.26/2.73	0.85/2.77	0.39/1.11
Proposed	D-iSAM2	**0.33/1.20**	0.19/0.76	**0.47**/1.37	0.33/1.54	**0.18/1.15**	0.59/2.35	**0.30/1.07**
Proposed	RISE [31]	**0.32/1.22**	**0.13/0.36**	0.59/1.64	0.32/1.52	0.32/2.97	0.59/2.35	**0.31/1.06**
Proposed	iSAM2	0.70/2.27	0.35/0.94	0.77/2.40	0.62/2.64	0.32/3.46	0.60/2.95	0.50/1.78
Huber	iSAM2	1.00/5.56	0.51/2.29	0.84/1.90	0.64/2.34	0.57/7.17	0.79/3.14	0.61/1.79

TABLE II

3D ABSOLUTE POSITION ERROR (METERS) USING 1HZ RANGE MEASUREMENTS (AVERAGE/MAX). IF THE AVERAGE ERROR IS GREATER THAN 10 METERS WE CLASSIFY THE RESULT AS FAILURE.

Noise model	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7
Gaussian	1.81/8.49	1.07/2.99	0.68/3.30	0.73/2.42	2.90/15.67	**0.91/3.97**	1.06/3.68
Huber	Fail	Fail	Fail	Fail	1.03/5.73	Fail	1.07/2.35
Proposed	**0.49/1.44**	**0.45/1.61**	**0.61/3.01**	**0.73/2.28**	**0.64/4.15**	1.30/5.54	**0.46/1.37**

A. Evaluation of proposed ranging model: 40Hz

We evaluate our proposed asymmetric model against Gaussian, also standard m-Estimators Cauchy and Huber, with the proposed D-iSAM2 estimator. The resulting metrics are shown in Table I. Our implicit hybrid method with asymmetric model beats all baselines in most runs, in both average and maximum trajectory errors.

B. Evaluation of proposed ranging model: lower frequency

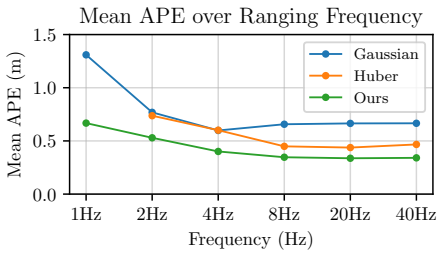

Fig. 7. Mean 3D APE over all datasets with different ranging frequency. Huber fails to give results on 1Hz (high APE due to estimation divergence).

Reducing the ranging frequency will enable localization with lower power, at the cost of increased difficulty of outlier identification because open-loop IMU integration's accuracy degrades quickly over time. We still use D-iSAM2 as incremental estimator, and keep all other parameters same as previous section.

In Table II we pushed our evaluation further to only use 1Hz ranging measurements. We find using Huber model causes complete estimation divergence on most of the runs, since Huber loss function identifies the true line-of-sight measurements as NLOS measurements. The Gaussian model, effectively treating all measurements as inliers, does not fail, but also leads to higher trajectory error. In Figure 7 we show the trajectory error results on a range of frequencies between 1Hz to 40Hz. We find that our proposed model always has the best trajectory accuracy at all ranging frequencies, and Huber is the second best except for its divergence observed at 1Hz (hence not plotted).

C. Evaluation of proposed D-iSAM2

TABLE III

PER ISAM2 UPDATE RUNTIME COMPARISON BETWEEN ESTIMATORS.

Estimator	Average Runtime (ms)	Max Runtime (ms)
iSAM2	21.33	83.19
D-iSAM2	19.21	79.46
RISE [31]	66.85	201.21

We also compare our proposed D-iSAM2 estimator against the RISE [31] Dogleg optimizer. All parameters remain the same except the type of optimizer used, and ranging frequency is set to 40Hz. Initial delta for the Dogleg algorithm is set at 0.1. The runtime of each estimator is shown in Table III while the quality of the optimized solution is presented in the lower part of Table I. Huber loss + iSAM2 is used as a baseline. Observe that the solution quality of RISE in some sequences matches those obtained by our algorithm, however, RISE does so consuming a significantly longer runtime. D-iSAM2 reaches similar runtime as vanilla iSAM2, while beating iSAM2 on trajectory accuracy to a large extent.

VII. CONCLUSION

We demonstrate that our proposed asymmetric noise model handles real-world range noises better than existing m-Estimators, with ranging measurement rate as low as 1Hz, which paves the way forward to accurate, energy-efficient indoor navigation using wireless ranging. We also show that our novel trust-region based incremental solver effectively handles the nonlinear range-IMU fusion problem. In most cases, we achieve both better convergence and lower run time than RISE and vanilla iSAM2. Finally, we show an end-to-end system which can localize accurately in indoor spaces using only IMU and UWB ranging measurements, with the hardware and software system we built, which has great potential in future low-power wearable devices.

REFERENCES

[1] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*. IEEE, 2007, pp. 3565–3572.

[2] S. Leutenegger, P. Furgale, V. Rabaud, M. Chli, K. Konolige, and R. Siegwart, "Keyframe-based visual-inertial slam using nonlinear optimization," *Proceedings of Robotis Science and Systems (RSS) 2013*, 2013.

[3] T. Qin, P. Li, and S. Shen, "VINS-Mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.

[4] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM: A versatile and accurate monocular SLAM system," *IEEE transactions on robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.

[5] T. Polonelli, S. Schlapfer, and M. Magno, "Performance Comparison between Decawave DW1000 and DW3000 in low-power double side ranging applications," in *2022 IEEE Sensors Applications Symposium (SAS)*. Sundsvall, Sweden: IEEE, Aug. 2022, pp. 1–6.

[6] K. Jiokeng, G. Jakllari, A. Tchana, and A.-L. Beylot, "When FTM Discovered MUSIC: Accurate WiFi-based Ranging in the Presence of Multipath," in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*. Toronto, ON, Canada: IEEE, Jul. 2020, pp. 1857–1866.

[7] M. Zhao, T. Chang, A. Arun, R. Ayyalasomayajula, C. Zhang, and D. Bharadia, "ULoc: Low-Power, Scalable and cm-Accurate UWB-Tag Localization and Tracking for Indoor Applications," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 3, pp. 1–31, Sep. 2021.

[8] J. Roderick, H. Krishnaswamy, K. Newton, and H. Hashemi, "Silicon-Based Ultra-Wideband Beam-Forming," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 8, pp. 1726–1739, Aug. 2006.

[9] Z. Qin, Z. Meng, Z. Li, N. Gao, Z. Zhang, Q. Meng, and D. Zhen, "Compensating the NLoS Occlusion Errors of UWB for Pedestrian Localization With MIMU," *IEEE Sensors Journal*, vol. 23, no. 11, pp. 12 146–12 158, Jun. 2023.

[10] J. A. Corrales, F. A. Candelas, and F. Torres, "Hybrid tracking of human operators using IMU/UWB data fusion by a Kalman filter," in *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*. Amsterdam The Netherlands: ACM, Mar. 2008, pp. 193–200.

[11] C. C. Cossette, M. Shalaby, D. Saussie, J. R. Forbes, and J. Le Ny, "Relative Position Estimation Between Two UWB Devices With IMUs," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4313–4320, Jul. 2021.

[12] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard, and F. Dellaert, "iSAM2: Incremental smoothing and mapping using the Bayes tree," *The International Journal of Robotics Research*, vol. 31, no. 2, pp. 216–235, Feb. 2012.

[13] V. Tran, Z. Dai, N. Trigoni, and A. Markham, "DeepCIR: Insights into CIR-based Data-driven UWB Error Mitigation," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Kyoto, Japan: IEEE, Oct. 2022, pp. 13 300–13 307.

[14] M. Kolakowski and J. Modelski, "First path component power based NLOS mitigation in UWB positioning system," in *2017 25th Telecommunication Forum (TELFOR)*. Belgrade: IEEE, Nov. 2017, pp. 1–4.

[15] Y. Cao, A. Dhekne, and M. Ammar, "6fit-a-part: A protocol for physical distancing on a custom wearable device," in *2020 IEEE 28th International Conference on Network Protocols (ICNP)*. IEEE, 2020, pp. 1–12.

[16] Y. Kong, C. Li, Z. Chen, and X. Zhao, "Recognition of Blocking Categories for UWB Positioning in Complex Indoor Environment," *Sensors*, vol. 20, no. 15, p. 4178, Jul. 2020.

[17] T. Lupton and S. Sukkarieh, "Visual-Inertial-Aided Navigation for High-Dynamic Motion in Built Environments Without Initial Conditions," *IEEE Trans. Robotics*, vol. 28, no. 1, pp. 61–76, Feb. 2012.

[18] E. Strömberg, "Smoothing and mapping of an unmanned aerial vehicle using ultra-wideband sensors," Master's thesis, KTH, Automatic Control / KTH, Automatic Control, 2017.

[19] F. Jiang, V. Agrawal, R. Buchanan, M. Fallon, and F. Dellaert, "iMHS: An Incremental Multi-Hypothesis Smoother," *arXiv:2103.13178 [cs]*, Mar. 2021.

[20] Q. Huang, C. Pu, D. Fourie, K. Khosoussi, J. P. How, and J. J. Leonard, "NF-iSAM: Incremental Smoothing and Mapping via Normalizing Flows," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. Xi'an, China: IEEE, May 2021, pp. 1095–1102.

[21] M. Hsiao and M. Kaess, "MH-iSAM2: Multi-hypothesis iSAM using Bayes Tree and Hypo-tree," in *2019 International Conference on Robotics and Automation (ICRA)*. Montreal, QC, Canada: IEEE, May 2019, pp. 1274–1280.

[22] K. J. Doherty, Z. Lu, K. Singh, and J. J. Leonard, "Discrete-Continuous Smoothing and Mapping," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 12 395–12 402, Oct. 2022.

[23] G. H. Lee, F. Fraundorfer, and M. Pollefeys, "Robust pose-graph loop-closures with expectation-maximization," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Tokyo: IEEE, Nov. 2013, pp. 556–563.

[24] V. D. Barnett, "Evaluation of the Maximum-Likelihood Estimator where the Likelihood Equation has Multiple Roots," *Biometrika*, vol. 53, no. 1/2, p. 151, Jun. 1966.

[25] Z. Zhang, "Parameter estimation techniques: A tutorial with application to conic fitting," *Image and Vision Computing*, vol. 15, no. 1, pp. 59–76, Jan. 1997.

[26] J. R. Carpenter and C. N. D'Souza, "Navigation Filter Best Practices," NASA, Tech. Rep., 2018.

[27] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation," in *Robotics: Science and Systems (RSS)*, 2015.

[28] M. Schuh, "Accurate UWB-based Localization using Online Calibrated and Corrected Single-Sided Two-Way Ranging," Ph.D. dissertation, TU Graz, May 2019.

[29] D. M. Rosen, M. Kaess, and J. J. Leonard, "Robust Incremental Online Inference Over Sparse Factor Graphs: Beyond the Gaussian Case," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, Karlsruhe, Germany, May 2013.

[30] J. Huai, Y. Lin, Y. Zhuang, and M. Shi, "Consistent Right-Invariant Fixed-Lag Smoother with Application to Visual Inertial SLAM," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 7, pp. 6084–6092, May 2021.

[31] D. M. Rosen, M. Kaess, and J. J. Leonard, "RISE: An Incremental Trust-Region Method for Robust Online Sparse Least-Squares Estimation," *IEEE Trans. Robotics*, vol. 30, no. 5, pp. 1091–1108, Oct. 2014.

[32] D. McGann, J. G. Rogers, and M. Kaess, "Robust Incremental Smoothing and Mapping (riSAM)," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. London, United Kingdom: IEEE, May 2023, pp. 4157–4163.

[33] J. Nocedal and S. J. Wright, *Numerical Optimization*, ser. Springer Series in Operations Research and Financial Engineering. Springer New York, 2006.

[34] K. Somasundaram, J. Dong, H. Tang, J. Straub, M. Yan, M. Goesele, J. J. Engel, R. De Nardi, and R. Newcombe, "Project aria: A new tool for egocentric multi-modal AI research," *arXiv preprint arXiv:2308.13561*, 2023.

[35] F. Dellaert and M. Kaess, *Factor Graphs for Robot Perception*. Foundations and Trends in Robotics, Vol. 6, 2017.

[36] M. Grupp, "Evo: Python package for the evaluation of odometry and SLAM." 2017.