2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

Towards Real-Time Non-Gaussian SLAM for Underdetermined Navigation

Dehann Fourie[1], Nicholas R. Rypkema[1,2], Pedro Vaz Teixeira[1,2], Sam Claassens[3], Erin Fischell[2], and John Leonard[1]

***Abstract*—This paper presents a method for processing sparse, non-Gaussian multimodal data in a simultaneous localization and mapping (SLAM) framework using factor graphs. Our approach demonstrates the feasibility of using a *sum-product* inference strategy to recover functional belief marginals from highly non-Gaussian situations, relaxing the prolific unimodal Gaussian assumption. The method is more focused than conventional multi-hypothesis approaches, but still captures dominant modes via multi-modality. The proposed algorithm exists in a trade space that spans the anticipated uncertainty of measurement data, task-specific performance, sensor quality, and computational cost. This work leverages several major algorithm design constructs, including *clique recycling*, to put an upper bound on the allowable computational expense – a major challenge in non-parametric methods. To better demonstrate robustness, experimental results show the feasibility of the method on at least two of four major sources of non-Gaussian behavior: i) the first introduces a canonical range-only problem which is always underdetermined although composed exclusively from Gaussian measurements; ii) a real-world AUV dataset, demonstrating how ambiguous acoustic correlator measurements are directly incorporated into a non-Gaussian SLAM solution, while using *dead reckon tethering* to overcome short term computational requirements.**

I. Introduction

Many robotic navigation applications require state-estimation methods beyond unimodal solutions due to four identified sources of non-Gaussian/multimodal behavior: i) non-linearity in measurement models [1]; ii) uncertain data association [2], [3]; iii) underdetermined problems (i.e. more unknowns than constraints) [4]; iv) physical measurement process is inherently ambiguous. This paper investigates latter two of these four identified sources, with the goal of developing robust simultaneous localization and mapping (SLAM). In addition, this paper addresses two major problems facing many perception systems, namely: handling of non-Gaussian data in an easy-to-understand factor graph framework, and methodologies that enable real-time navigation solutions. This work aims to demonstrate the feasibility, complexity, and importance of non-Gaussian solutions/methods in the context of applications dealing with such multimodal and highly uncertain problems.

[1]D. Fourie, N. R. Rypkema, P. V. Teixeira, J. Leonard are with the MIT, Cambridge, MA 02139, USA. {fourie, rypkema, pvt}@mit.edu

[2] Woods Hole Oceanographic Institution, Woods Hole, MA, USA. efischell@whoi.edu

[3]S. Claassens is with University of Chicago, IL, USA. samueldc@uchicago.edu

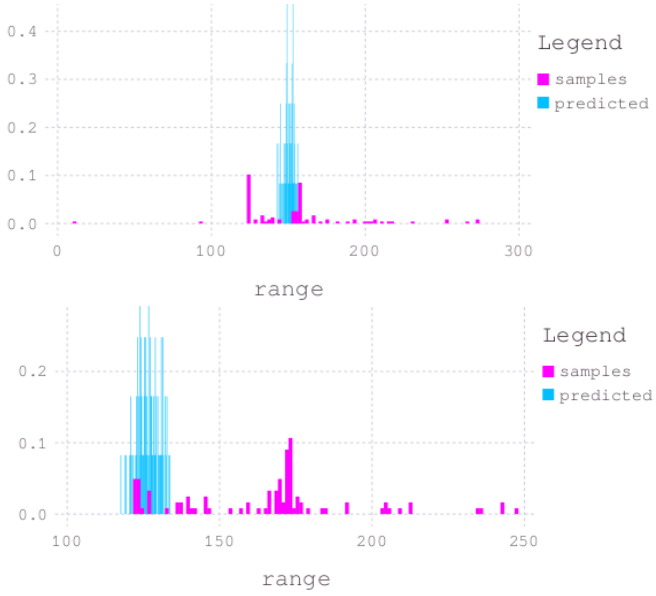

Fig. 1: Predicted and measured (sampled) acoustic range probability density values from an underwater robotic data set. Top: the predicted measurement after initialization coincides with a secondary, incorrect mode; after joint inference, the predicted range coincides with the correct mode. Bottom: after inference, the predicted range matches a smaller but correct measurement mode.

A major concern with non-Gaussian/multimodal work is the computational complexity that results from relaxing the ubiquitous Gaussian parametric assumption; the proposed method presents a adjustable upper bounded computational load approach during operation of the system, while resolving fully continuous non-Gaussian marginal estimates for variables of interest. Our method leverages several concepts from both existing state-estimation as well as novel researched methods, and casts these ideas in a natural, common, and easy to understand methodology. This paper is complementary to, as well as heavily dependent on, prior computational reduction by characterizing marginalization operations on the underlying Bayes tree [5]. The associated *clique recycling* methods are briefly discussed here in Section III-B and forms a critical part of the development towards real-time non-Gaussian SLAM. Our claim is that the proposed approach is viable for real-time applications given an order of magnitude improvement in computational performance (either by improved implementation or hardware).

Contributions of this work include: i) non-Gaussian/multimodal factor graph [6] data-fusion approach that naturally handles multiple modes; ii) incorporation of

dead reckon tethers, alongside *clique recycling* [5], as a path towards real-time belief based non-Gaussian SLAM; iii) introduction of an easily reproducible canonical test problem for underdetermined (i.e. non-Gaussian multimodal) SLAM; iv) demonstration of a single range beacon non-Gaussian SLAM navigation solution on real data; and simultaneous v) demonstration of an end-to-end (sensors-to-solution) sum-product solution without requiring any Gaussian assumptions in the navigation processing pipeline.

The next section discusses related methods and prior solutions, followed by our proposed approach and experimental results. Two sets of results are presented, illustrating cases iii) and iv) of the four sources of non-Gaussian behavior: a) a novel canonical underdetermined (i.e. singular) system and b) real-world AUV results for single-beacon navigation looking at multimodal processing with ambiguity in range measurement. Lastly, interpretations and a conclusion is given.

II. Related Work

In many acoustic and radio navigation applications (e.g. underwater acoustic or remote sensor system positioning), the range to a source or target is estimated through correlation of the received signal with a copy of the transmitted signal (i.e. matched filter [7]). The response is a energy intensity over time-of-arrival which does not always produce a clear distinct unimodal peak, see Vaz Teixeira et. al. [8]. Bayesian (e.g. Kalman) filtering techniques are often employed to track the *correct* portion of the correlator output, however, most filtering methods marginalized-out all but the current state making them ill-suited [9] for further development of joint smoothing with other multi-sensor data. Although SLAM is often used for multi-sensor processing, direct use of this correlator intensity output is poorly suited to state-of-the-art Gaussian-only methods, as both acoustic and electromagnetic signals are subject to phenomena (e.g. non-linearities, multi-path propagation, interference) that result in a highly non-Gaussian, and often multimodal correlator output. Figure 1 illustrates multipath interference in the correlator output of a real-world acoustic time-of-flight signal, prioritized selection of the highest peaks would, in this case, be incorrect.

Recent work by Rypkema et. al. [10] demonstrates the feasibility of a passive on-vehicle beam-forming and particle filtering approach, using a low-cost Bluefin SandShark AUV [11]. Particle filtering followed by SLAM [12] means that measurement data association decisions are marginalized out and the new Gaussian-only measurement likelihood assumption is made for consumption by the SLAM framework. Seminal work by Newman and Leonard [13] demonstrated a smoothing approach by applying pre-filtered LBL range measurements in an EKF-SLAM framework. This work was further extended by Olson et. al. [14] by improving the method of outlier rejection.

Leitinger et. al. [15] propose a range-based SLAM technique that explicitly handles multipath effects by instantiating both real and virtual (image) beacons. This requires handling association of range measurements, which is performed using a probabilistic method. Pacholska et al. [16] tackle the range-only localization problem by employing a parametric model of the vehicle trajectory, and formulate the recovery as a quadratic optimization problem, dependent on the trajectory parameters and range measurements, whose uncertainty is assumed to follow a zero-mean Gaussian. This problem is then relaxed to a linear form, and conditions are presented for the successful recovery of trajectories.

Fourie et. al. [2] recently introduced a framework for performing non-parametric inference over a factor graph, providing a method of solving graphs that include non-Gaussian and multimodal constraints. This framework was used in [4] to estimate the trajectory of an AUV using raw acoustic measurements output by matched filtering and beamforming, without having to integrate a complex and carefully tuned outlier rejection pre-filter. This approach allowed for direct processing of non-Gaussian acoustic measurements produced by conventional beamforming and matched filtering (as shown in Figure 1), without having to artificially 'fit' Gaussians to these distributions. In this work, we seek to improve prior navigation work [10] by directly converting (i.e. tightly coupling) the acoustic front-end correlator as a probability likelihood [17] into a non-Gaussian factor graph SLAM framework, thereby avoiding any maximum point selection or Gaussian belief assumptions in the sensing to inference pipeline.

III. Method

For many applications, the long term robustness of the stand-alone navigation system is significantly more important than the instantaneous position accuracy, as long as post-hoc data processing can recover the desired accuracy and up to some tolerable risk. For example, long-endurance AUVs require a reliable navgition solution, which can leverage all ambiguous sensor data including uncertain loop closures in a sparse feature environment. Navigation reliability may be improved for AUV applications either with higher accuracy (and more expensive/power-hungry) sensors, or by improved data fusion of heterogeneous sensor data streams. The approach presented in this paper trades instantaneous accuracy in favor of more navigation robustness using methods in non-Gaussian/multimodal SLAM that fundamentally address the underlying probability uncertainties at an increased computational load. We posit that—considering the AUV use-case example—a slower SLAM solution cycle time (e.g. on the order of tens of seconds) for a full factor graph solution is viable, if the benefits of significantly increased robustness helps enable long-term navigation autonomy. Furthermore, SLAM-based joint data post-processing can also allows provide accuracy improvements for data products frequently required for survey, inspection, and mapping applications.

A. Factor Graph SLAM

While exploring the world, a robot incrementally collects sets of measurements $\mathbf{Z} = \{\boldsymbol{z}_{1:t}\}$ at each time step t. These measurements are made locally involving only a few hidden variable states. This sparse structure is best leveraged using a factor graph modeling language, as illustrated in Figure 2.

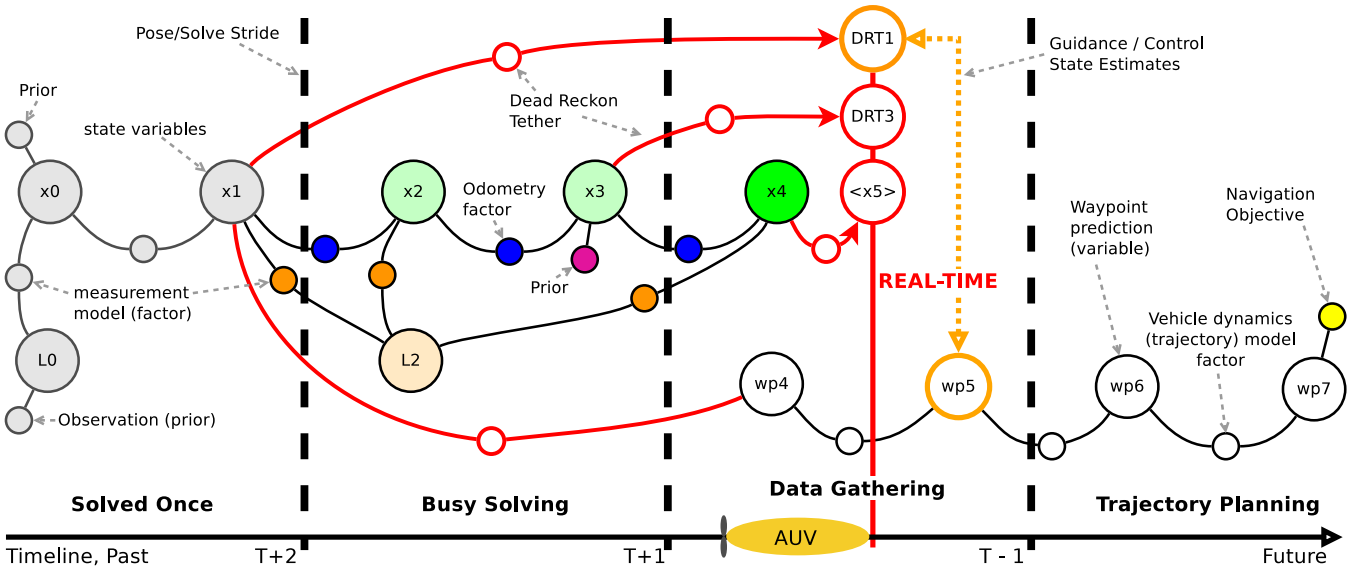

Fig. 2: Illustration of factor graph construction during AUV vehicle collecting data from the environment: odometry factors (blue) relate the successive vehicle poses (green) and range constraints (brown) relate the vehicle poses to the beacon location (light brown). Dead reckon tethers (red) provide an unified means of near instantaneous state-estimation while a new SLAM is busy solving, while incorporating trajectory optimization.

Each sensor measurement is modeled by a their own measurement function $\phi_i(\boldsymbol{\Theta}_i, \boldsymbol{z}_i)$ over a subset of variables $\boldsymbol{\Theta}_i$. These functions either construct or already are a probabilistic likelihood model, or factor, $p_{\phi_i} = p(\boldsymbol{Z}_{\phi_i} = \boldsymbol{z}_{\phi_i} | \boldsymbol{\Theta}_i) \in \mathcal{P}$; $\mathcal{P}$ denotes the space of all allowable probability density functions, and the union of all likelihoods is called the set of factors $\mathcal{F} = \bigcup_i p_{\phi_i}$.

Assuming that measurements $\boldsymbol{z}_i, \boldsymbol{z}_j$ are taken from statistically independent processes $\boldsymbol{Z}_i, \boldsymbol{Z}_j$ given the factor model $\{\phi_{i \cup j}\} (i \neq j)$, the inference problem can be formulated as a joint probability distribution through a large product operation. Considering the variables of interest $\boldsymbol{\Theta}$, factors $\mathcal{F}$, and a collection of edges $\mathcal{E}$ between them, the factor graph $\mathcal{G} = \{\mathcal{F}, \boldsymbol{\Theta}, \mathcal{E}\}$ encodes

$$p(\boldsymbol{\Theta} | \mathbf{Z}) \propto \prod_i p(\boldsymbol{z}_{\phi_i} | \boldsymbol{\Theta}_i) \prod_j p(\boldsymbol{\Theta}_j) \tag{1}$$

where partition scaling taken as constant due to the independence assumption. This allows further simplification of the unnormalized joint probability function. By the chain rule, this product of independent measurements $\mathbf{Z}$—through likelihoods $p(\boldsymbol{z}_{\phi_i} | \boldsymbol{\Theta}_i)$ and variable priors $p(\boldsymbol{\Theta}_j)$—represents the unnormalized, non-Gaussian posterior joint probability density [6]. The inference task now becomes to invert the system and estimate the belief over state variables given the data $p(\mathbf{Z} | \boldsymbol{\Theta}) \rightarrow p(\boldsymbol{\Theta} | \mathbf{Z}) \in \mathcal{P}$ that likely produced the received measurements $\mathbf{Z}$.

B. Solving Chapman-Kolmogorov (& Clique Recycling)

While cyclic factor graphs are well suited for combining various sensory data streams, their cyclic structure (i.e. loop-closures) renders them hard to compute algebraically. Tractability is addressed by factorization into an acyclic *Bayes tree* [18], [5]: the process first selects a variable order [19] to guide construction of a chordal *Bayes network* [20] using a *bipartite elimination game* [21], and then discovers *cliques* through the *maximum cardinality search* [22]. Each clique k in the represents a partial joint over a subset of *frontal* variables $\Theta_{F,k}$ with their connected factors likelihoods i' and priors j'

$$p(\Theta_{F,k} | \Theta_{S,k}, \mathbf{z}_k) \propto \prod_{i'}^{C_k} p(\mathbf{z}_{i'} | \Theta_{i'}) \prod_{j'}^{C_k} p(\Theta_{j'}). \tag{2}$$

These factors separate the frontal variables from the rest of the factor graph via the *separator* variables $\Theta_{S,k}$, with $\Theta_{C,k} = \Theta_{F,k} \cup \Theta_{S,k}$. The tree of cliques describes a sequence of statistical dependencies staring from the leaves along parallel branches up to the root clique:

$$M_{k|\mathbf{Y}}(\Theta_{C,k}) \propto p(\Theta_{F,k} | \Theta_{S,k}, \mathbf{z}_k) \prod_u m_{u|\mathbf{Y}}(\Theta_{S,u}), \tag{3}$$

where incoming belief messages from each child clique u is the out-marginalized partial joint posterior of separators:

$$m_{u|\mathbf{Y}}(\Theta_{S,u}) = \int_{\Xi} M_{u|\mathbf{Y}}(\Theta_{C,u}) \, d\theta_{F,u}, \tag{4}$$

where variable marginalizations take place over their respective manifold domains [23], $\theta_{F,k} \in \Xi_k$.

Eqs. (3) and (4) describe the general Chapman-Kolmogorov transit integrals that both describe familiar belief propagation as well as Bayesian filtering. The filtering case can be recognized when considering an acyclic factor graph forming a long chain—a hidden Markov model—and chronological variable order for elimination into a linear Bayes tree; thus the messages $m_{u|Y}(\Theta_{S,u})$ represent the forward propagating belief that is familiar in recursive methods such as particle or Kalman filtering (assuming messages $m_{k|\mathbf{Y}}$ are Gaussian).

It is precisely this structure of the Bayes tree that allows the computational complexity to be controlled by the designer, by selecting how many of the clique beliefs $M_{k|\mathbf{Y}}(\Theta_{C,k})$ should be recalculated; leaving the (likely older) cliques as though

out-marginalized – this is directly analogous to conventional filtering. The characteristics of marginalization and recycling operation on the Bayes tree are described in more detail by Fourie et. al. [5]. The criteria by which variables (ultimately cliques) in the factor graph are marked for marginalization can vary; for the purposes of this work, a first-in-first-out fixed-lag window of non-marginalized variables is used. As the mission progresses, more variables and in essence a larger part of the Bayes tree is recycled, thereby focusing computational resources only on the newest variables added to the graph.

Lastly, the numerical calculation of clique partial joint posteriors $M_{k|\boldsymbol{Y}}\left(\Theta_{C,k}\right)$ can be performed by any chosen method. Traditionally, when conguency of parametric Gaussian-only assumptions is used, the clique's *maximum-a-posteriori* estimate is sought from the dense least-squares objective $\theta^*_{C,k} = \arg\min_{\theta_{C,k}} - \log M_{k|\boldsymbol{Y}}\left(\Theta_{C,k}\right)$. Adaptations such as max-mixtures at factor [24] or clique level multi-hypothesis [25] may also be used to solve for eqs. (3) and (4).

However, a more capable non-Gaussian approach is need to resolve the Chapman-Kolmogorov equations generated in this paper, since the resulting joint partial posteriors are simultaneously highly non-Gaussian and multimodal. Our approach follows earlier work on non-parametric belief propagation [2], [23] to resolve full functional belief estimates of eq. (4), which together with clique recycling on the Bayes tree is called the multimodal incremental smoothing and mapping (MM-iSAM) algorithm [4]. The Kernel Density [26] posterior estimation approach was inspired by [27] and is best summarized as a multiscale, nested-mini-batch Gibbs sampling process which is based on successive functional convolution and function product operations.

C. Dead Reckon Tether

A meaningful fixed-lag window still requires non-zero computation time which likely exceeds the instantaneous real-time timing requirements. The *dead reckon tethers (DRT)*, in Figure 2, provides high-rate localization estimates by using a mutable odometry factor (either parametric or non-parametric) starting from the latest *gauge variable* (i.e. recently SLAM-inferred). The tethers operate directly on the factor graph and can use the same odometry factor mechanization as the ongoing SLAM solution, but are separately updated from sensor data at high-rate to produce near instantaneous state estimates. By spanning over an unsolved section of the factor graph, the DRT allows for additional SLAM compute time. This architecture also provides a framework for pricing various sensor technologies: the cost of a blind inertial odometry solution [4], [28], for example, should sustain the required drift accuracies over the computation period required, see Figure 2.

After a large data fusion calculation, the step transition from DRT1 to DRT3 as primary estimate may cause a discrete jump in the state estimate, which may induce poor/unsatisfactory control performance in a conventional closed-loop system architecture. However, DRTs also provide a solution by combining trajectory planning and model predictive control optimization routines into the factor graph as a joint inference process, as suggested by Dong et al. [29] and Fourie [4]. In Figure 2, DRT1 is anchored on gauge-variable x_1 and the result can be directly compared to waypoint target wp_5 for a real-time control solution. Forthcoming work using intelligent message passing on the Bayes tree will address how to prevent errors in the future trajectory portion from adversely effecting the historic portions of the factor graph, a problem identified in [29].

IV. Experiments and Results

Two experiments (using simulated and real data) demonstrate how the proposed method copes with two of the four identified sources of non-Gaussian behavior described in Section I: iii) underdetermined problem situations and iv) physical measurement process uncertainties. The experiments aim to demonstrate the method feasibility and that the method is possibly one of the first truly non-Gaussian-belief (sum-product) SLAM navigation approaches with controllable computational complexity. The first experiment investigates underdetermined SLAM without the use of DRT, and using only Gaussian factors that still produce highly non-Gaussian posteriors. The second example is based on real-world data and demonstrates a single range system with the addition of non-Gaussian factors and DRT tethers for instantaneous solutions. All results presented here were performed using the MM-iSAM implementation [30], running on a laptop equipped with an i7-9750H CPU at 2.6GHz and 64GB of RAM.

A. Underdetermined Range-Only Example

To aid with the lack of canonical non-Gaussian SLAM problems in literature, the following simulated two dimensional problem is constructed to highlight how underdetermined behavior manifests, and that it cannot be directly solved with conventional linear algebra techniques. A progression of repeated solutions follows the addition of each new pose variable, along with sparse measurements so that the number of unknown dimensions always exceeds the number of measurement dimensions. In effect, this problem is a singular system for which a unimodal inversion $p\left(\mathbf{Z} \mid \Theta\right) \rightarrow p\left(\Theta \mid \mathbf{Z}\right)$ is not possible, and the resulting variable marginal posteriors are non-Gaussian and multimodal.

Figure 3 shows the ground truth of the vehicle (cyan line) as it moves between four unique landmark beacons in a series of 13 distinct positions tracing out the shape of the letter 'e'. The true locations of the two prior known red landmarks are L1:$(10, 30)$, L2:$(30 - 30)$, and two prior unknown cyan landmarks are L3:$(80, 40)$ and L4:$(120, -50)$. Range measurements from pose to beacon are indicated in Plots A and B using magenta lines and do not exceed $150\ m$. The noteworthy vehicle positions are at P1:$(0, 0)$, P3:$(100, 0)$, P5:$(100, 100)$, P7:$(0, 100)$, P8:$(0, 50)$, P9:$(0, -50)$, P10:$(0, -100)$, P12:$(100, -100)$, P13:$(100, -50)$, with intermittent values midway between their neighbors. All odometry and beacon ranges are taken as 1-dimensional, normally distributed values only

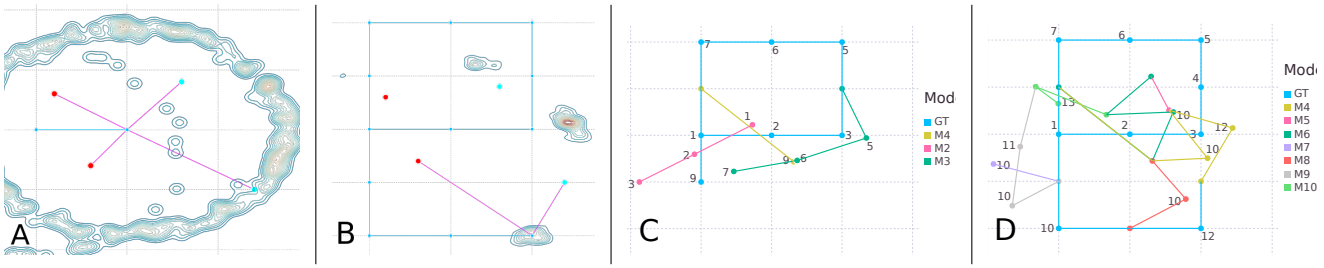

Fig. 3: Underdetermined trilateration example. *A*. Marginal posterior belief for landmark L_4 in bottom right with nearly infinite possible distinct locations. *B*. Four distinct modes are shown for pose 12 when relying on two landmarks. *C*. Trajectory hypotheses for first 9 poses - note the two starting location modes, M1 (cyan, coincident with grount truth GT) and M2 (pink). *D*. Various trajectory hypotheses for pose 13.

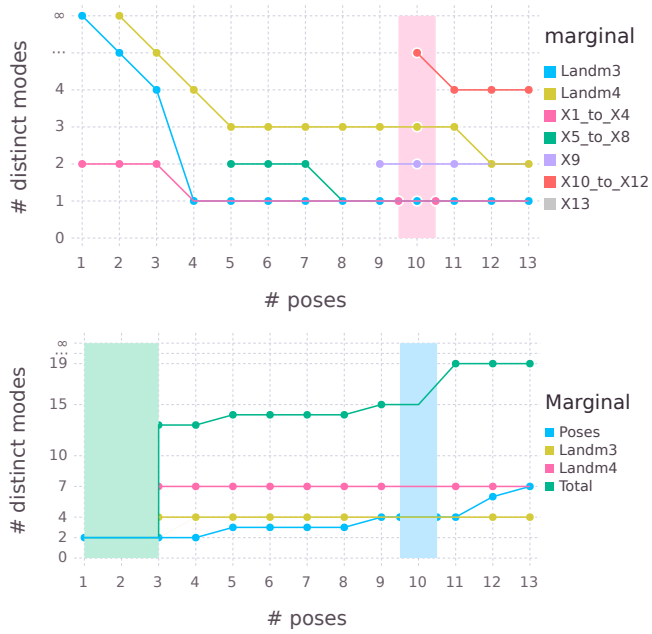

Fig. 4: *Top*, distinct sustained modes in all pose and landmark marginals. Similar variables are grouped by color throughout the entire robot trajectory. For example, landmark L_4 is only observed at the second pose with near infinite possible distinct locations. Notice how the mode counts collapse as more measurement likelihood information is gathered. Due to weaker constraint, L_4 maintains around three distinct modes throughout the latter part of the robot trajectory. *Bottom*, total number of distinct modes considered in the system.

– e.g. the odometry factor from P1 to P2 is taken as a $\mathcal{N}(\mu = 50, \sigma = 3)$. After each new pose and ranging, a full factor graph solution is computed, using the previous values as initialization for the next.

After adding P2 and solving, the contour lines in Plot A Figure 3 shows the marginal posterior belief for the unknown landmark (bottom right) $p\left(\hat{L}_3 \,|\, \mathbf{Z}\right)$, clearly showing the highly non-Gaussian and multimodal nature of the beliefs. Note that marginal posteriors for P1, P2, L1, L2, L3, and L4 were all simultaneously recovered, but Plot A only shows the marginal for L3. Plot B shows the marginal over pose $p\left(\hat{P}_{12} \,|\, \mathbf{Z}\right)$, which is clearly has four distinct visible modes, as well as significant probability density at the true location (bottom right of the 'e'). Note that the ground truth and main trajectory shown in Plot B are coincident as the cyan lines, and that the true mode is not the maximum-a-posteriori point.

To simplify the visualization, Plots C and D depict a manual interpretation of the variable posterior marginals by fitting normal densities to each of the distinct modes found by MM-iSAM. Since the starting position is initially unknown, the two ranges from red landmarks L1 and L2 produce two intersections – i.e two modes for possible starting location. This is shown in plot C of Figure 3 by the cyan (M1) and magenta (M2) lines for position 1. P1 through P3 sustain two modes (M1 and M2) post inference, and pose 4 drops down to only one mode. Plot C shows four trajectory hypotheses recovered from data for the first nine possible pose locations, and Plot D indicates the resolved trajectory possibilities through pose P13. Each of the traces are alternate trajectory hypotheses that was recovered by the multimodal rempresentation, and which fit the data in a sufficiently likely way. Positions 5 through 7 again see a uptick in modes, shown as the cyan and green (M3) traces, and similar for poses 8 and 9 with the cyan and yellow trace (M4). The process continues with modes M5 through M10 through to pose 13, as shown in the right plot of Figure 3.

Note that the main mode of interest, M1, is tracked throughout the entire SLAM problem growth and completely solves after each new pose addition. The parasitic loss of modes are not permanent and can be reacquired depending on stochastics, solution quality, or introduction of new data, see Figures 3 and 4. There still is a low likelihood second mode at P4 as the total mirror image around the two red landmarks, but this mode is dropped due to its stochastic low likelihood. Repeat runs sometimes produces the same set of "nearby" modes surrounding the M2 case, while dropping the M1 case presented here. While the error statistics of this example have less immediate value for the scope of discussion, and given a lack of meaningful metric for non-Gaussian distributions, a long format discussion with more detail is available for further reference [4]. Furthermore, runnable source code for this example is available as open-source software along with the MM-iSAM solver [30].

B. Underwater Acoustic Ranging Navigation

This experiment aims to show, by means of post-processing real-world AUV data, the feasibility of developing a fully non-Gaussian factor graph SLAM solution. To demonstrate the benefits of the non-Gaussian SLAM approach, a solution is constructed using only a subset of the available data—i.e. a single hydrophone/beacon—from the five-hydrophone data that was previously independently shown by Rypkema et. al. [12], [10]. This demonstrates a

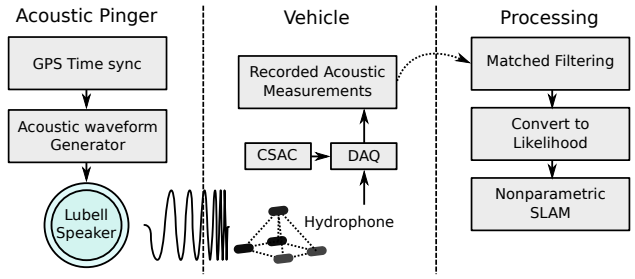

Fig. 5: Diagrammatic system overview. A single acoustic beacon periodically transmits an acoustic signal which is synchronously received by a five element iUSBL hydrophone array mounted on the AUV. This recorded acoustic data is processed offline by matched filtering and beamforming, followed by sampling, and finally by the non-parametric SLAM framework.

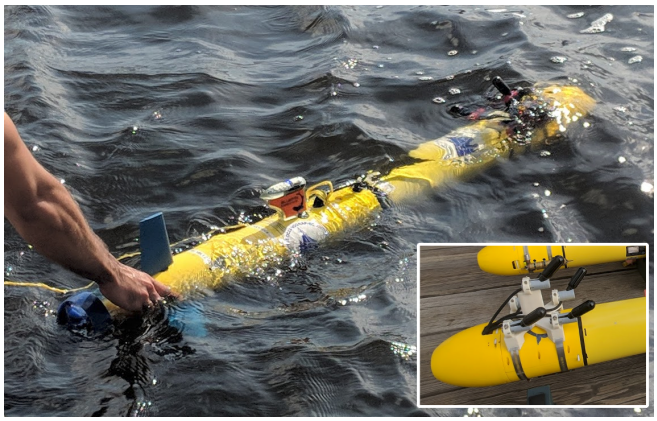

Fig. 6: Photo of our commercial Bluefin SandShark miniature AUV outfitted with a custom pyramidal iUSBL hydrophone array mounted above the nose (inset).

real-world underdetermined measurement (i.e. weak observability) environment, while simultaneously using inherently ambiguous (non-Gaussian and multimodal) raw acoustic correlator measurements, see examples in Figure 1. This increased ambiguity is either very difficult or impossible to resolve with conventional unimodal Gaussian-only approaches. Lastly, this experiment is also complementary to related work on synthetic aperture sonar SLAM [1].

1) Experiment Setup: A Bluefin SandShark AUV [11] outfitted with our custom acoustic payload [12], shown in Figure 6, was operated on a section of the Charles River adjacent to the MIT sailing pavilion. The acoustic beacon was submerged and fastened to the pavilion dock in a fixed, known location at a depth of approximately $1\ m$, transmitting a unique GPS time-synchronized acoustic chirp every second. Reference data was recorded by means of a second time-synchronized, but frequency separated, beacon also mounted to the dock and post-processed to provide a conventional reference long baseline (LBL) $\arg\max$ solution for the vehicle position during the mission. The AUV was preprogrammed to travel back-and-forth along the dock in a racetrack pattern of length $80\ m$ and width $10\ m$ at a depth of $2\ m$ and a speed of $1.4\ m/s$. The duration of the mission was set to $1400\ s$. After processing the reference data, we found the actual track covered to be stretched longer and racetrack orbits slowly drifted South-East.

2) Matched Filter Ranging and SLAM: The acoustic chirps were received via an AUV-mounted hydrophone, digitized and recorded by custom onboard electronics. The AUV is also equipped with a GPS time-synchronized chip scale atomic clock (CSAC) to allow for one-way ranging over reasonable mission times. One-way time-of-flight is calculated as the energy intensity response [7] from the recorded acoustic data using matched filtering (MF) and a template of the chirp template signal

$$y[n] = \sum_{k=0}^{N-1} h[n-k]x_i[k], \tag{5}$$

where $x[n]$ is the signal recorded by the hydrophone, and $h[n]$ is the template of the known transmitted waveform that produces the energy intensity response $y[n]$. This energy intensity over the time-of-flight response is converted into a *pseudo range probability likelihood density* by deconvolving the best known channel model $m[n]$ and exponentiating the negative energy response function [17] $p\left(\, Z_\rho \,|\, \boldsymbol{X}_i, \boldsymbol{L} \,\right)$. The probability likelihood $z \sim Z_\rho$ is normalized and converted into the range-domain $z \in R$ by multiplying samples by the speed of sound c and dividing by the sampling rate S – examples are shown in Figure 1.

In contrast to parametric methods, we seek the full energy over time-of-flight response. Natively constructing range factors from acoustic correlator likelihood $p\left(\, Z_\rho \,|\, \boldsymbol{X}_i, \boldsymbol{L} \,\right)$ avoids an maximum point selection, and therefore any loss of information only occurs during joint SLAM inference. Thereby preserving the entire range of possible values. Odometry factors are also added to the factor graph using a more familiar multivariate normal error model from a heading and speed motion models using inertial measurement unit (IMU) gyro rotation and propeller rotation rate. Uniformly distributed ($\pm 5^o$) magnetometer-derived yaw priors are also added to each pose. A location prior is supplied to the beacon as well as to the first AUV pose in the odometry chain using

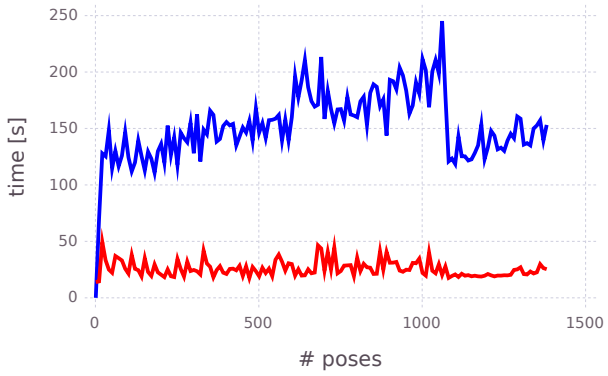

Fig. 7: AUV example: SLAM solve times for individual full non-Gaussian belief factor graph solutions using MM-iSAM – a new solution occurs after every ten new poses are added. The red trace is the initialization time s for all variables that have been added since completion of the last solve. The blue trace is the Bayes tree solution time with clique recycling [5] enabled and a lag-window of 50 unmarginalized variables selected.

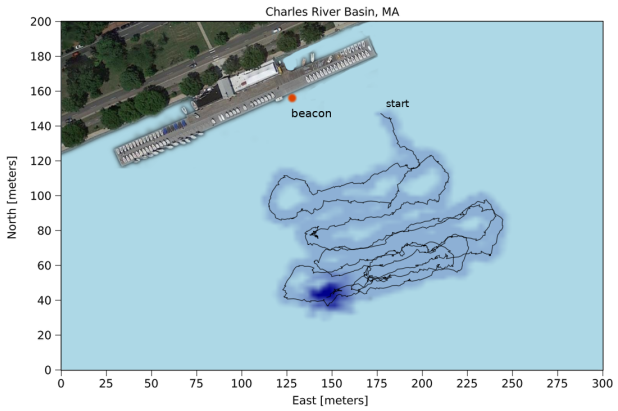

Fig. 8: SLAM result showing AUV performing several passes off the Dock at the MIT sailing pavilion. The black trace shows the mean point estimate for each pose in the trajectory, surrounded by a shaded region representing the actual marginal beliefs of each pose superimposed on a common image.

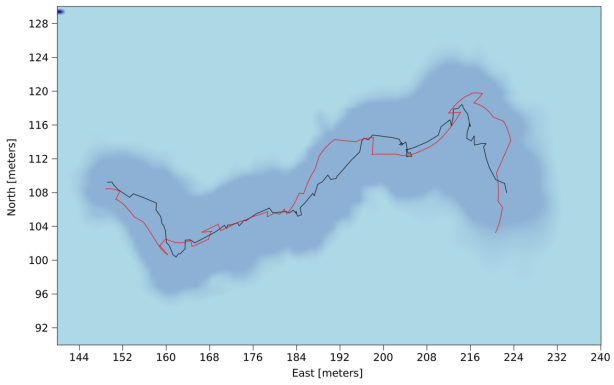

Fig. 9: Zoomed in region of Figure 8 comparing the "real-time" dead reckoning tether DRT1 (red trace) to the non-Gaussian SLAM solution estimate (black), which is available after the joint solution is calculated. Note although the tether result is noisy, it is within the marginal beliefs for each of the poses.

GPS just before the vehicle dives. Figure 9 shows a portion from Figure 8 to better illustrate *Busy Solving* and *Data Gathering* aspects from Figure 2.

C. Single Beacon Results

The results utilize both the clique recycling [5] and DRT, Section III-C, methodologies from Section III to upper bound on computational load. Figure 8 shows the MIT sailing pavilion and resolved trajectory. The red trace represents the active DRT solution (DRT1 in Figure 2) which is available near instantaneously during the *Data Gathering* phase, but is more noisy. The black trace in Figures 8 and 2 shows a maximum-a-posteriori point on the full variable marginal beliefs (i.e. equivalent to maximum-a-posteriori), which become numerically available a short time after real-time as variables from the *Busy Solving* segment are completed. As discussed in Section III-B, the shaded region in Figure 8 is from variables that have been (at least) *Solved Once*. The dark blue shading is the current belief estimate of the latest pose, and the lighter shading is the marginal belief probability mass (i.e. confidences) of older poses.

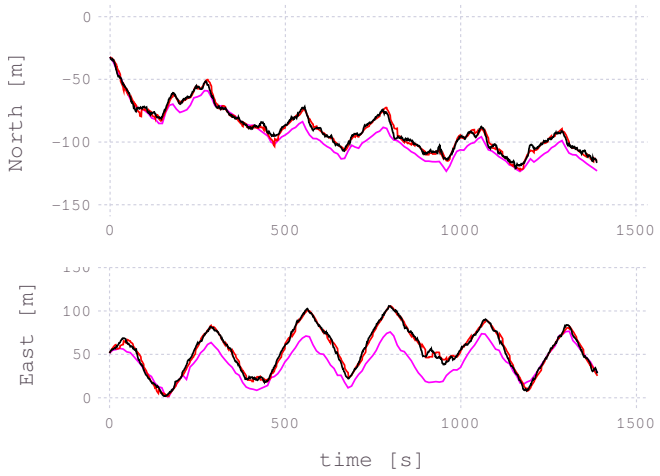

Fig. 10: North and East position time series estimate. The SLAM result is shown as the black trace, with DRT as red and two beacon LBL reference in magenta.

Both the most recent instantaneous DRT position estimate (red trace) and SLAM-derived maximum-a-posteriori position estimate (black trace) are compared to the LBL reference estimate (magenta trace) in Figure 10. The SLAM solution shown in Figure 10 does not exactly correspond to the LBL solution during the center part of the trajectory, but returns during the last quarter. This is due the trajectory running tangential to the range measurement towards the single beacon, where displacement along the direction of the dock is weakly observable. The discrepancy in range measurements at this point is small compared to the size of East-West offset.

The quantified 1 σ variation between the SLAM solution and the dead reckon tether is less than 4 m over the entire 20 min trajectory, as shown in Figure 11. Furthermore, the distribution shapes are concentrated around zero which is indicates minimal systematic errors in the dead reckon tether approach. In addition, the aggregate errors between SLAM and the LBL reference are tabulated in I for various parameter sweeps relevant to the proposed approach, such as marginalization lag-window size or odometry factor noise level.

Lastly, the computation time of SLAM with tree clique recycling enabled is indeed bounded, as shown in Figure 7. The result also shows that for a lag-window of 50 un-

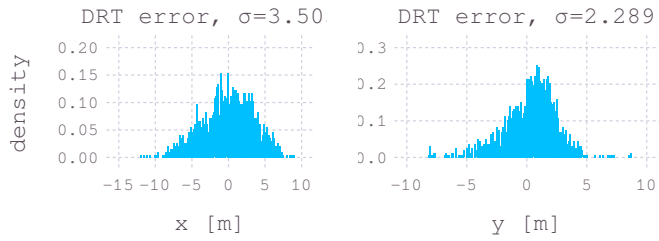

Fig. 11: Error between instantaneous DRT solution and delayed full SLAM result aggregated for all poses.

TABLE I: Summarized AUV 1-beacon error results as 1 σ from post-hoc 2-beacon LBL reference, as pose by error similar to Figure 10. The pose-solve stride is fixed at every 10 poses, but range factor ρ stride either 3 or 4, tree-based marginalization *lag* length, and odometry noise (small/medium/large).

	$^{ps}/_{\rho}$	lag	Q_{odo}	$1\sigma\,[m]$	$RMSE\,[m]$
#1	3	50	Med.	9.77	19.6
#2	3	50	Lrg.	4.24	12.0
#3	4	30	Sml.	4.49	9.52
#4	4	30	sml.	4.28	9.41
#5	4	30	Lrg.	42.75	22.32

marginalized variables, the total initialization and SLAM processing time is, on average, limited to less than 200 s per solve iteration, sustained throughout the entire 1400 s trajectory. As far as we are aware, this is the first time a fully non-Gaussian/multimodal SLAM approach of this kind has been presented.

V. Conclusion

While the results presented herein were computed with prototype level software, we posit that with a 10-fold improvement in processing performance—mostly by means of improved software performance and resource utilization—non-Gaussian SLAM approaches should become viable for real-time applications in the years ahead. This work demonstrates a factor graph centric approach for multi-sensor data fusion is viable and can operate in circumstances where conventional methods fail, specifically in cases of underdetermined constraint or high measurement ambiguity. Thus, front-end navigation system design efforts can be simplified since ambiguous measurements no longer have to be processed/filtered to fit a unimodal Gaussian assumption. The improved flexibility of such an approach does incur a higher computational cost, however, this paper in combination with other work [5] outlined a mechanism to upper bound the computational load of the non-Gaussian/multimodal SLAM solution.

Acknowledgements

This work was partially supported by the Office of Naval Research under grants N00014-18-1-2832 and MURI N00014-19-1-2571.

References

[1] M. Y. Cheung, D. Fourie, N. R. Rypkema, P. V. Teixeira, H. Schmidt, and J. Leonard, "Non-gaussian slam utilizing synthetic aperture sonar," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 3457–3463.

[2] D. Fourie, J. Leonard, and M. Kaess, "A nonparametric belief solution to the Bayes tree," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems, IROS*, Daejeon, Korea, Oct 2016.

[3] K. Doherty, D. Fourie, and J. Leonard, "Multimodal semantic slam with probabilistic data association," in *2019 international conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 2419–2425.

[4] D. Fourie, "Multi-Modal and Inertial Sensor Solutions to Navigation-type Factor Graphs," Ph.D. dissertation, Massachusetts Institute of Technology and Woods Hole Oceanographic Institution, 2017.

[5] D. Fourie, A. T. Espinoza, M. Kaess, and J. J. Leonard, "Characterizing marginalization and incremental operations on the Bayes tree," in *International Workshop on Algorithmic Foundations of Robotics (WAFR)*, Finalnd, June 2020.

[6] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, Feb. 2001.

[7] H. L. Trees, *Detection, Estimation, and Modulation Theory: Part III. Radar-Sonar Signal Processing and Gaussian Signal in Noise*. John Wiley & Sons, Incorporated, 2001.

[8] P. V. Teixeira, M. Kaess, F. S. Hover, and J. J. Leonard, "Multibeam data processing for underwater mapping," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, Oct. 2018, pp. 1877–1884.

[9] F. Dellaert, "Square Root SAM: Simultaneous location and mapping via square root information smoothing," in *Robotics: Science and Systems (RSS)*, Cambridge, MA, 2005.

[10] N. R. Rypkema, E. M. Fischel, and H. Schmidt, "Closed-loop single-beacon passive acoustic navigation for low-cost autonomous underwater vehicles," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 641–648.

[11] A. Underwood and C. Murphy, "Design of a micro-auv for autonomy development and multi-vehicle systems," in *OCEANS 2017-Aberdeen*. IEEE, 2017, pp. 1–6.

[12] N. R. Rypkema, E. M. Fischell, and H. Schmidt, "One-way travel-time inverted ultra-short baseline localization for low-cost autonomous underwater vehicles," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, 2017, pp. pre–print.

[13] P. Newman and J. Leonard, "Pure range-only sub-sea SLAM," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, vol. 2, Sep. 2003, pp. 1921–1926.

[14] E. Olson, J. Leonard, and S. Teller, "Fast iterative alignment of pose graphs with poor initial estimates," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, May 2006, pp. 2262–2269.

[15] E. Leitinger, F. Meyer, F. Hlawatsch, K. Witrisal, F. Tufvesson, and M. Z. Win, "A belief propagation algorithm for multipath-based SLAM," *IEEE transactions on wireless communications*, vol. 18, no. 12, pp. 5613–5629, 2019.

[16] M. Pacholska, F. Duembgen, and A. Scholefield, "Relax and recover: Guaranteed range-only continuous localization," *IEEE Robotics and Automation Letters*, 2020.

[17] P. M. Woodward, *Probability and Information Theory, with Applications to Radar: International Series of Monographs on Electronics and Instrumentation*. Elsevier, 2014, vol. 3.

[18] M. Kaess, V. Ila, R. Roberts, and F. Dellaert, "The Bayes tree: An algorithmic foundation for probabilistic robot mapping," in *Intl. Workshop on the Algorithmic Foundations of Robotics, WAFR*, Singapore, Dec. 2010.

[19] T. Davis, J. Gilbert, S. Larimore, and E. Ng, "A column approximate minimum degree ordering algorithm," *ACM Trans. Math. Softw.*, vol. 30, no. 3, pp. 353–376, 2004.

[20] J. Pearl, "Bayesian networks," *Department of Statistics, UCLA*, 2011.

[21] P. Heggernes and P. Matstoms, *Finding good column orderings for sparse QR factorization*. University of Linköping, Department of Mathematics, 1996.

[22] R. E. Tarjan and M. Yannakakis, "Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs," *SIAM Journal on computing*, vol. 13, no. 3, pp. 566–579, 1984.

[23] D. Fourie, P. V. Teixeira, and J. Leonard, "Non-parametric mixed-manifold products using multiscale kernel densities," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nov 2019, pp. 6656–6662.

[24] E. Olson and P. Agarwal, "Inference on networks of mixtures for robust robot mapping," *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 826–840, 2013.

[25] 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)*. IEEE, 2019, pp. 1274–1280.

[26] B. W. Silverman, *Density estimation for statistics and data analysis*. CRC press, 1986, vol. 26.

[27] E. Sudderth, A. Ihler, M. Isard, W. Freeman, and A. Willsky, "Nonparametric belief propagation," *Communications of the ACM*, vol. 53, no. 10, pp. 95–103, 2010.

[28] D. Fourie and J. Leonard, "Inertial odometry with retroactive sensor calibration," Jun. 11 2019, uS Patent 10,317,214.

[29] J. Dong, M. Mukadam, F. Dellaert, and B. Boots, "Motion planning as probabilistic inference using gaussian processes and factor graphs." in *Robotics: Science and Systems*, vol. 12, 2016.

[30] Contributors and Packages. (2020) Caesar.jl. https://github.com/JuliaRobotics/Caesar.jl.