Semantic Distributed Multi-Robot Classification, Localization, and Mapping With a Viewpoint Dependent Classifier Model
Supplementary Material

Vladimir Tchuiev and Vadim Indelman

This document provides supplementary material to [1]. Therefore, it should not be considered a self-contained document, but instead regarded as an appendix of [1]. Throughout this report, all notations and definitions are with compliance to the ones presented in [1].

1 Supplementary Derivation

We present a relation that is used in equation derivation; Let A be a random variable conditioned on the set $\{B_i\}$ of random variables B_i that are independent from each other. By using Bayes Law, we can split the conditional probability $\mathbb{P}(A|\{B_i\})$ to a product of conditional probabilities:

$$\mathbb{P}(A|\{B_i\}) = \frac{\mathbb{P}(\{B_i\}|A)\mathbb{P}(A)}{\mathbb{P}(\{B_i\})} = \frac{\prod_i \mathbb{P}(B_i|A)}{\prod_i \mathbb{P}(B_i)}\mathbb{P}(A). \tag{1}$$

Using Bayes Law again on each element in the product, we reach the following expression:

$$\mathbb{P}(A|\{B_i\}) = \prod_i \left(\frac{\mathbb{P}(A|B_i)}{\mathbb{P}(A)}\right)\mathbb{P}(A). \tag{2}$$

This allow to express a random variable as a multiplication of conditionals, which will be useful to separate local and external measurements.

2 Derivation of $\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R, \mathcal{H}_k^{R-})$

Recall splitting $\mathcal{H}_k^{R-}$ into prior history $\mathcal{H}_{k-1}^R$ and non-local measurements & actions:

$$\mathcal{H}_k^{R-} = \mathcal{H}_{k-1}^R \cup \Delta\mathcal{H}_k^{R-}. \tag{3}$$

We then use the above definition and relation (2) to split $\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R, \mathcal{H}_k^{R-})$ into a product of two beliefs, one that depends on prior history, and one that depends on external new measurements:

$$\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R, \mathcal{H}_k^{R-}) = \mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R) l \frac{\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R, \Delta\mathcal{H}_k^{R-})}{\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R)} \frac{\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R, \mathcal{H}_{k-1}^R)}{\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R)}. \tag{4}$$

This formulation allows us to isolate the new information sent by other robots at time k, from information already used for inference at previous times. Next, we have to address that not all known objects are present in the sent local beliefs. Because the priors are assumed independent between poses and classes, $\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R) = \mathbb{P}(\mathcal{X}_k^R \backslash x_k^r)$. From $\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R, \mathcal{H}_{k-1}^R)$ we can split $\mathcal{X}_k^R \backslash x_k^r$ into poses of objects that are involved in $\mathcal{H}_{k-1}^R$ and ones that do not as:

$$\mathbb{P}(\mathcal{X}_k^R \backslash x_k^r | C_k^R, \mathcal{H}_{k-1}^R) = \mathbb{P}(\mathcal{X}_{\text{new},k}^{o,R} | C_{\text{new},k}^R, \mathcal{H}_{k-1}^R)\mathbb{P}(\mathcal{X}_{k-1}^R | C_k^R, \mathcal{H}_{k-1}^R). \tag{5}$$

The authors are with the Department of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel. {vovatch, vadim.indelman}@technion.ac.il.

Poses of objects that r wasn't aware of at time $k-1$ are independent of $\mathcal{H}^R_{k-1}$, and without measurements, $\mathcal{X}^{o,R}_{\text{new},k}$ are independent of $C^R_{\text{new},k}$ as well. In addition, $\mathcal{X}^R_{k-1}$ is independent of classes of objects that are observed only at time k, thus we can write:

$$\mathbb{P}(\mathcal{X}^R_k \backslash x^r_k | C^R_k, \mathcal{H}^R_{k-1}) = \mathbb{P}(\mathcal{X}^{o,R}_{\text{new},k}) \cdot \mathbb{P}(\mathcal{X}^R_{k-1} | C^R_{k-1}, \mathcal{H}^R_{k-1}), \tag{6}$$

which is the prior for poses of newly known objects at time step k, multiplied by the conditional continuous belief for objects already known. Similarly to Eq. (6), the prior for $\mathcal{X}^R_k \backslash x^r_k$ is separated to previously known and new objects:

$$\mathbb{P}(\mathcal{X}^R_k \backslash x^r_k) = \mathbb{P}(\mathcal{X}^{o,R}_{\text{new},k}) \mathbb{P}(\mathcal{X}^R_{k-1}), \tag{7}$$

therefore we can write:

$$\frac{\mathbb{P}(\mathcal{X}^R_k \backslash x^r_k | C^R_k, \mathcal{H}^R_{k-1})}{\mathbb{P}(\mathcal{X}^R_k \backslash x^r_k | C^R_k)} = \frac{b^R_{k-1}}{\mathbb{P}(\mathcal{X}^R_{k-1})}, \tag{8}$$

and substitute it into the rightmost fracture in Eq. (4). Then we cancel out $\mathbb{P}(\mathcal{X}^R_k \backslash x^r_k | C^R_k)$ and proceed to remove r's poses from the prior by:

$$\frac{\mathbb{P}(\mathcal{X}^R_k \backslash x^r_k | C^R_k, \Delta\mathcal{H}^{R-}_k)}{\mathbb{P}(\mathcal{X}^R_{k-1})} = \frac{\mathbb{P}(\mathcal{X}^{o,R}_k | C^R_k, \Delta\mathcal{H}^{R-}_k)}{\mathbb{P}(\mathcal{X}^{o,R}_{k-1})}, \tag{9}$$

as robot r's poses up until time $k-1$ are independent from the new external measurements. Finally, after factoring out $\mathbb{P}(\mathcal{X}^R_k \backslash x^r_k | C^R_k)$, and Eq. (8) and (9) with Eq. (4) we reach the following expression that is used in the paper:

$$\mathbb{P}(\mathcal{X}^R_k \backslash x^r_k | C^R_k, \mathcal{H}^{R-}_k) = b^R_{k-1} \cdot \frac{\mathbb{P}(\mathcal{X}^{o,R}_k | C^{o,R}_k, \Delta\mathcal{H}^{R-}_k)}{\mathbb{P}(\mathcal{X}^{o,R}_{k-1})} \tag{10}$$

3 Dependency Between Object Classes

In this section we show that the classes of two objects are not independent. We present a simple example where c_1 and c_2 be the underlying classes of objects 1 and 2 respectively. Let $\mathcal{H}^R$ be the total measurement history, including semantic measurements z^{sem}_1 and z^{sem}_2 for objects 1 and 2 respectively. Recall that measurements are assumed independent from each other. Using the Bayes Law:

$$\mathbb{P}(c_1, c_2 | \mathcal{H}^R) \propto \mathbb{P}(c_1, c_2 | \mathcal{H}^R \backslash z^{sem}_1, z^{sem}_2) \mathbb{P}(z^{sem}_1 | c_1) \mathbb{P}(z^{sem}_2 | c_2). \tag{11}$$

We use a viewpoint dependent classifier model, so we must marginalize $\mathbb{P}(z^{sem}_1 | c_1)\mathbb{P}(z^{sem}_2 | c_2)$ by the corresponding relative viewpoints, denoted x^{rel}_1 and x^{rel}_2 respectively:

$$\mathbb{P}(z^{sem}_1 | c_1)\mathbb{P}(z^{sem}_2 | c_2) = \int_{x^{rel}_1, x^{rel}_2} \mathbb{P}(z^{sem}_1 | c_1, x^{rel}_1)\mathbb{P}(z^{sem}_2 | c_2, x^{rel}_2)\mathbb{P}(x^{rel}_1, x^{rel}_2 | \mathcal{H}^R) dx^{rel}_1 dx^{rel}_2. \tag{12}$$

From the above equation, the condition for c_1 and c_2 to be independent is that x^{rel}_1 and x^{rel}_2 must be independent, which is not true in the general case, thus c_1 and c_2 are dependent.

4 Derivation of $\frac{\mathbb{P}(\mathcal{X}^{o,R}_k | C^R_k, \Delta\mathcal{H}^{R-}_k)}{\mathbb{P}(\mathcal{X}^{o,R}_k)}$ (Continuous Belief Update)

Using the relation (2) we can split the blue part by separating the new measurements and actions per robot, excluding r itself as it is not present in $\Delta\mathcal{H}^{R-}_k$:

$$\mathbb{P}(\mathcal{X}^{o,R}_k | C^R_k, \Delta\mathcal{H}^{R-}_k) = \prod_{k', r' \in R \backslash r} \left(\frac{\mathbb{P}(\mathcal{X}^{o,R}_{k'} | C^{r'}_{k'}, \Delta\mathcal{H}^{r'}_{k'})}{\mathbb{P}(\mathcal{X}^{o,R}_{k'})} \right) \mathbb{P}(\mathcal{X}^{o,R}_k). \tag{13}$$

From that, we will address every element in the product. Poses of objects that r' doesn't observe locally can be canceled out as follows, leaving only the object poses that r' observed:

$$\frac{\mathbb{P}(\mathcal{X}^{o,R}_{k'} | C^{r'}_{k'}, \Delta\mathcal{H}^{r'}_{k'})}{\mathbb{P}(\mathcal{X}^{o,R}_{k'})} = \frac{\mathbb{P}(\mathcal{X}^{o,r'}_{k'} | C^{r'}_{k'}, \Delta\mathcal{H}^{r'}_{k'})}{\mathbb{P}(\mathcal{X}^{o,r'}_{k'})}. \tag{14}$$

Then, using relation (2) again, we can expand $\mathbb{P}(\mathcal{X}_{k'}^{o,r'}|C_{k'}^{r'},\mathcal{H}_{k'}^{r'})$ to separate between known prior and new measurements:

$$\mathbb{P}(\mathcal{X}_{k'}^{o,r'}|C_{k'}^{r'},\mathcal{H}_{k'}^{r'}) = \mathbb{P}(\mathcal{X}_{k'}^{o,r'})\frac{\mathbb{P}(\mathcal{X}_{k'}^{o,r'}|C_{k'}^{r'},\mathcal{H}_{k'-l'}^{r'})}{\mathbb{P}(\mathcal{X}_{k'}^{o,r'})}\frac{\mathbb{P}(\mathcal{X}_{k'}^{o,r'}|C_{k'}^{r'},\Delta\mathcal{H}_{k'}^{r'})}{\mathbb{P}(\mathcal{X}_{k'}^{o,r'})}, \tag{15}$$

with l' being the time difference between subsequent slots of r (that can be 0 if the slot isn't updated). Then we take out all the poses that aren't dependent on the prior information, and we reach the definition of $\xi_{k-1}^{r,r'}$, i.e. the marginal object poses at the previous time.

$$\frac{\mathbb{P}(\mathcal{X}_{k'}^{o,r'}|C_{k'}^{r'},\mathcal{H}_{k'-l'}^{r'})}{\mathbb{P}(\mathcal{X}_{k'}^{o,r'})} = \frac{\mathbb{P}(\mathcal{X}_{k'-l'}^{o,r'}|C_{k'-l'}^{r'},\mathcal{H}_{k'-l'}^{r'})}{\mathbb{P}(\mathcal{X}_{k'-l'}^{o,r'})} \doteq \xi_{k-1}^{r,r'}. \tag{16}$$

With the definition of $\xi_{k}^{r,r'}$, and by substituting Eq. (16) into Eq. (15) we reach the expression for a single element of the product in Eq. (13):

$$\frac{\mathbb{P}(\mathcal{X}_{k'}^{o,r'}|C_{k'}^{r'},\Delta\mathcal{H}_{k'}^{r'})}{\mathbb{P}(\mathcal{X}_{k'}^{o,r'})} = \frac{\xi_{k}^{r,r'}}{\xi_{k-1}^{r,r'}}. \tag{17}$$

Finally, substituting Eq. (17) we reach the expression for the external continuous update belief:

$$\frac{\mathbb{P}(\mathcal{X}_{k}^{o,R}|C_{k}^{R},\Delta\mathcal{H}_{k}^{R-})}{\mathbb{P}(\mathcal{X}_{k}^{o,R})} = \prod_{r'\in R}\frac{\xi_{k}^{r,r'}}{\xi_{k-1}^{r,r'}}. \tag{18}$$

5 Derivation of $\frac{\mathbb{P}(C_k^R|\Delta\mathcal{H}_k^{R-})}{\mathbb{P}(C_k^R)}$ (Discrete Belief Update)

The discrete belief update is similar to the continuous in its derivation, with probability over class realization, rather than object poses. Again, using the relation (2) we can split the red part by separating the new measurements and actions per robot, excluding r itself as it is not present in $\Delta\mathcal{H}_k^{R-}$:

$$\mathbb{P}(C_k^R|\Delta\mathcal{H}_k^{R-}) = \prod_{k',r'\in R\setminus r}\left(\frac{\mathbb{P}(C_{k'}^R|\Delta\mathcal{H}_{k'}^{r'})}{\mathbb{P}(C_{k'}^R)}\right)\mathbb{P}(C_{k'}^R) \tag{19}$$

From that, we will address every element in the product. Classes of objects that r' doesn't observe locally can be canceled out as follows, leaving only the classes of object that r' observed:

$$\frac{\mathbb{P}(C_{k'}^R|\Delta\mathcal{H}_{k'}^{r'})}{\mathbb{P}(C_{k'}^R)} = \frac{\mathbb{P}(C_{k'}^{r'}|\Delta\mathcal{H}_{k'}^{r'})}{\mathbb{P}(C_{k'}^{r'})}. \tag{20}$$

Then, using relation (2) again, we can expand $\mathbb{P}(C_{k'}^{r'}|\mathcal{H}_{k'}^{r'})$ to separate between known prior and new measurements:

$$\mathbb{P}(C_{k'}^{r'}|\mathcal{H}_{k'}^{r'}) = \mathbb{P}(C_{k'}^{r'})\frac{\mathbb{P}(C_{k'}^{r'}|\mathcal{H}_{k'-l'}^{r'})}{\mathbb{P}(C_{k'}^{r'})}\frac{\mathbb{P}(C_{k'}^{r'}|\Delta\mathcal{H}_{k'}^{r'})}{\mathbb{P}(C_{k'}^{r'})}, \tag{21}$$

with l' being the time difference between subsequent slots of r (that can be 0 if the slot isn't updated). Then we take out all the classes of objects that not appear in prior information, and we reach the definition of $\phi_{k-1}^{r,r'}$, i.e. the marginal object poses at the previous time.

$$\frac{\mathbb{P}(C_{k'}^{r'}|\mathcal{H}_{k'-l'}^{r'})}{\mathbb{P}(C_{k'}^{r'})} = \frac{\mathbb{P}(C_{k'-l'}^{r'}|\mathcal{H}_{k'-l'}^{r'})}{\mathbb{P}(C_{k'-l'}^{r'})} \doteq \phi_{k-1}^{r,r'} \tag{22}$$

With the definition of $\phi_{k}^{r,r'}$, and by substituting Eq. (22) into Eq. (21) we reach the expression for a single element of the product in Eq. (19):

$$\frac{\mathbb{P}(C_{k'}^{r'}|\Delta\mathcal{H}_{k'}^{r'})}{\mathbb{P}(C_{k'}^{r'})} = \frac{\phi_{k}^{r,r'}}{\phi_{k-1}^{r,r'}} \tag{23}$$

Finally, substituting Eq. (23) we reach the expression for the external continuous update belief:

$$\frac{\mathbb{P}(C_{k}^{R}|\Delta\mathcal{H}_{k}^{R-})}{\mathbb{P}(C_{k}^{R})} = \prod_{r'\in R}\frac{\phi_{k}^{r,r'}}{\phi_{k-1}^{r,r'}}. \tag{24}$$

6 Simulation: Parameters

We consider a motion model with noise covariance $\Sigma_w = diag(0.003, 0.003, 0.001)$, and geometric model with noise covariance $\Sigma_v^{geo} = diag(0.1, 0.1, 0.01)$, both corresponding to position coordinates in meters and orientation in radians.

Our semantic model parameters are defined as:

$$\begin{aligned} h_c(c=1,\psi) &\doteq [0.25 \cdot \sin(\psi) + 0.75, 0.25(1-\sin(\psi))]^T \\ h_c(c=2,\psi) &\doteq [0.25(1-\sin(\psi)), 0.25 \cdot \sin(\psi) + 0.75]^T, \end{aligned}$$

where $h_c(c=i,\psi) \in \mathbb{R}^M$ is the predicted probability vector given object class c is i. Recall that our semantic measurements $z_k^{sem,r}$ are probability vectors as well. ψ is the relative orientation between robot and object, computed from the relative pose $x_k^{rel} \doteq x^o \ominus x_k$. The measurement covariance is defined via the square root information matrix, such that $\Sigma_c \doteq (R^T R)^{-1}$, and $R = \begin{bmatrix} 1.5 & -0.75 \\ 0 & 1.5 \end{bmatrix}$. Both the geometric and semantic measurements are limited to 10 meters from the robot's pose. The highest probability for ambiguous class measurements is at $\psi = -90°$, where $h_c = [0.5, 0.5]^T$ for both classes.

7 Simulation: Table of Stack Time Stamps

In this section we present a table of stack time stamps that indicates direct and indirect communication between robots in our scenario. Recall that the maximal communication radius is 10 meters, thus robots r_2 and r_3 communicate from time $k = 6$, robots r_1 starts communicating to others from time $k = 13$.

Time step	Stack of r_1	Stack of r_2	Stack of r_3
$k = 1$	t(r_1): 1 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 1
$k = 2$	t(r_1): 2 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 2
$k = 3$	t(r_1): 3 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 3 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 3
$k = 4$	t(r_1): 4 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 4 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 4
$k = 5$	t(r_1): 5 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 5 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 5
$k = 6$	t(r_1): 6 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 6 t(r_3): 5	t(r_1): 0 t(r_2): 5 t(r_3): 6
$k = 7$	t(r_1): 7 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 7 t(r_3): 6	t(r_1): 0 t(r_2): 6 t(r_3): 7
$k = 8$	t(r_1): 8 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 8 t(r_3): 7	t(r_1): 0 t(r_2): 7 t(r_3): 8
$k = 9$	t(r_1): 9 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 9 t(r_3): 8	t(r_1): 0 t(r_2): 8 t(r_3): 9
$k = 10$	t(r_1): 10 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 10 t(r_3): 9	t(r_1): 0 t(r_2): 9 t(r_3): 10
$k = 11$	t(r_1): 11 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 11 t(r_3): 10	t(r_1): 0 t(r_2): 10 t(r_3): 11
$k = 12$	t(r_1): 12 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 12 t(r_3): 11	t(r_1): 0 t(r_2): 11 t(r_3): 12
$k = 13$	t(r_1): 13 t(r_2): 12 t(r_3): 12	t(r_1): 12 t(r_2): 13 t(r_3): 12	t(r_1): 12 t(r_2): 12 t(r_3): 13
$k = 14$	t(r_1): 14 t(r_2): 13 t(r_3): 13	t(r_1): 13 t(r_2): 14 t(r_3): 13	t(r_1): 13 t(r_2): 13 t(r_3): 14
$k = 15$	t(r_1): 15 t(r_2): 14 t(r_3): 14	t(r_1): 14 t(r_2): 15 t(r_3): 14	t(r_1): 14 t(r_2): 14 t(r_3): 15

Time step	Stack of r_1	Stack of r_2	Stack of r_3
$k = 16$	t(r_1): 16 t(r_2): 15 t(r_3): 15	t(r_1): 15 t(r_2): 16 t(r_3): 15	t(r_1): 15 t(r_2): 15 t(r_3): 16
$k = 17$	t(r_1): 17 t(r_2): 16 t(r_3): 16	t(r_1): 16 t(r_2): 17 t(r_3): 16	t(r_1): 16 t(r_2): 16 t(r_3): 17
$k = 18$	t(r_1): 18 t(r_2): 17 t(r_3): 17	t(r_1): 17 t(r_2): 18 t(r_3): 17	t(r_1): 17 t(r_2): 17 t(r_3): 18
$k = 19$	t(r_1): 19 t(r_2): 18 t(r_3): 18	t(r_1): 18 t(r_2): 19 t(r_3): 18	t(r_1): 18 t(r_2): 18 t(r_3): 19
$k = 20$	t(r_1): 20 t(r_2): 19 t(r_3): 19	t(r_1): 19 t(r_2): 20 t(r_3): 19	t(r_1): 19 t(r_2): 19 t(r_3): 20
$k = 21$	t(r_1): 21 t(r_2): 20 t(r_3): 20	t(r_1): 20 t(r_2): 21 t(r_3): 20	t(r_1): 20 t(r_2): 20 t(r_3): 21
$k = 22$	t(r_1): 22 t(r_2): 21 t(r_3): 21	t(r_1): 21 t(r_2): 22 t(r_3): 21	t(r_1): 21 t(r_2): 21 t(r_3): 22
$k = 23$	t(r_1): 23 t(r_2): 22 t(r_3): 22	t(r_1): 22 t(r_2): 23 t(r_3): 22	t(r_1): 22 t(r_2): 22 t(r_3): 23
$k = 24$	t(r_1): 24 t(r_2): 23 t(r_3): 23	t(r_1): 23 t(r_2): 24 t(r_3): 23	t(r_1): 23 t(r_2): 23 t(r_3): 24
$k = 25$	t(r_1): 25 t(r_2): 24 t(r_3): 24	t(r_1): 24 t(r_2): 25 t(r_3): 24	t(r_1): 24 t(r_2): 24 t(r_3): 25
$k = 26$	t(r_1): 26 t(r_2): 25 t(r_3): 25	t(r_1): 25 t(r_2): 26 t(r_3): 25	t(r_1): 25 t(r_2): 25 t(r_3): 26
$k = 27$	t(r_1): 27 t(r_2): 26 t(r_3): 26	t(r_1): 26 t(r_2): 27 t(r_3): 26	t(r_1): 26 t(r_2): 26 t(r_3): 27
$k = 28$	t(r_1): 28 t(r_2): 27 t(r_3): 27	t(r_1): 27 t(r_2): 28 t(r_3): 27	t(r_1): 27 t(r_2): 27 t(r_3): 28
$k = 29$	t(r_1): 29 t(r_2): 28 t(r_3): 28	t(r_1): 28 t(r_2): 29 t(r_3): 28	t(r_1): 28 t(r_2): 28 t(r_3): 29
$k = 30$	t(r_1): 30 t(r_2): 29 t(r_3): 29	t(r_1): 29 t(r_2): 30 t(r_3): 29	t(r_1): 29 t(r_2): 29 t(r_3): 30

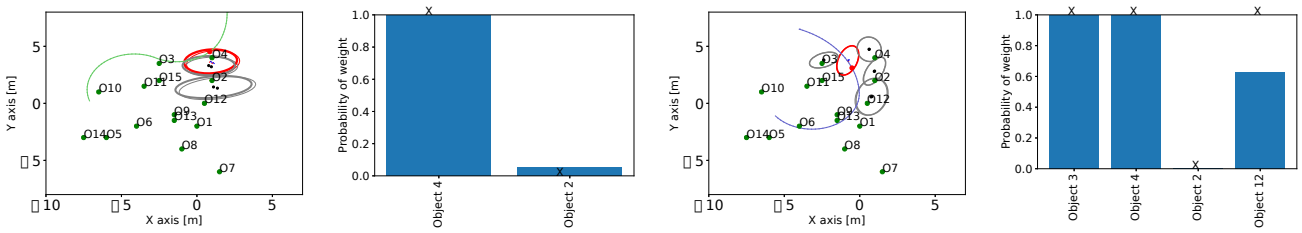
Time step	Stack of r_1	Stack of r_2	Stack of r_3
$k = 31$	t(r_1): 31 t(r_2): 30 t(r_3): 30	t(r_1): 30 t(r_2): 31 t(r_3): 30	t(r_1): 30 t(r_2): 30 t(r_3): 31
$k = 32$	t(r_1): 32 t(r_2): 31 t(r_3): 31	t(r_1): 31 t(r_2): 32 t(r_3): 31	t(r_1): 31 t(r_2): 31 t(r_3): 32
$k = 33$	t(r_1): 33 t(r_2): 32 t(r_3): 32	t(r_1): 32 t(r_2): 33 t(r_3): 32	t(r_1): 32 t(r_2): 32 t(r_3): 33
$k = 34$	t(r_1): 34 t(r_2): 33 t(r_3): 33	t(r_1): 33 t(r_2): 34 t(r_3): 33	t(r_1): 33 t(r_2): 33 t(r_3): 34
$k = 35$	t(r_1): 35 t(r_2): 34 t(r_3): 34	t(r_1): 34 t(r_2): 35 t(r_3): 34	t(r_1): 34 t(r_2): 34 t(r_3): 35
$k = 36$	t(r_1): 36 t(r_2): 35 t(r_3): 35	t(r_1): 35 t(r_2): 36 t(r_3): 35	t(r_1): 35 t(r_2): 35 t(r_3): 36
$k = 37$	t(r_1): 37 t(r_2): 36 t(r_3): 36	t(r_1): 36 t(r_2): 37 t(r_3): 36	t(r_1): 36 t(r_2): 36 t(r_3): 37
$k = 38$	t(r_1): 38 t(r_2): 37 t(r_3): 37	t(r_1): 37 t(r_2): 38 t(r_3): 37	t(r_1): 37 t(r_2): 37 t(r_3): 38
$k = 39$	t(r_1): 39 t(r_2): 38 t(r_3): 38	t(r_1): 38 t(r_2): 39 t(r_3): 38	t(r_1): 38 t(r_2): 38 t(r_3): 39
$k = 40$	t(r_1): 40 t(r_2): 39 t(r_3): 39	t(r_1): 39 t(r_2): 40 t(r_3): 39	t(r_1): 39 t(r_2): 39 t(r_3): 40
$k = 41$	t(r_1): 41 t(r_2): 40 t(r_3): 40	t(r_1): 40 t(r_2): 41 t(r_3): 40	t(r_1): 40 t(r_2): 40 t(r_3): 41
$k = 42$	t(r_1): 42 t(r_2): 41 t(r_3): 41	t(r_1): 41 t(r_2): 42 t(r_3): 41	t(r_1): 41 t(r_2): 41 t(r_3): 42
$k = 43$	t(r_1): 43 t(r_2): 42 t(r_3): 42	t(r_1): 42 t(r_2): 43 t(r_3): 42	t(r_1): 42 t(r_2): 42 t(r_3): 43
$k = 44$	t(r_1): 44 t(r_2): 43 t(r_3): 43	t(r_1): 43 t(r_2): 44 t(r_3): 43	t(r_1): 43 t(r_2): 43 t(r_3): 44
$k = 45$	t(r_1): 45 t(r_2): 44 t(r_3): 44	t(r_1): 44 t(r_2): 45 t(r_3): 44	t(r_1): 44 t(r_2): 44 t(r_3): 45

Time step	Stack of r_1	Stack of r_2	Stack of r_3
$k = 46$	t(r_1): 46 t(r_2): 45 t(r_3): 45	t(r_1): 45 t(r_2): 46 t(r_3): 45	t(r_1): 45 t(r_2): 45 t(r_3): 46
$k = 47$	t(r_1): 47 t(r_2): 46 t(r_3): 46	t(r_1): 46 t(r_2): 47 t(r_3): 46	t(r_1): 46 t(r_2): 46 t(r_3): 47
$k = 48$	t(r_1): 48 t(r_2): 47 t(r_3): 47	t(r_1): 47 t(r_2): 48 t(r_3): 47	t(r_1): 47 t(r_2): 47 t(r_3): 48
$k = 49$	t(r_1): 49 t(r_2): 48 t(r_3): 48	t(r_1): 48 t(r_2): 49 t(r_3): 48	t(r_1): 48 t(r_2): 48 t(r_3): 49
$k = 50$	t(r_1): 50 t(r_2): 49 t(r_3): 49	t(r_1): 49 t(r_2): 50 t(r_3): 49	t(r_1): 49 t(r_2): 49 t(r_3): 50
$k = 51$	t(r_1): 51 t(r_2): 50 t(r_3): 50	t(r_1): 50 t(r_2): 51 t(r_3): 50	t(r_1): 50 t(r_2): 50 t(r_3): 51
$k = 52$	t(r_1): 52 t(r_2): 51 t(r_3): 51	t(r_1): 51 t(r_2): 52 t(r_3): 51	t(r_1): 51 t(r_2): 51 t(r_3): 52
$k = 53$	t(r_1): 53 t(r_2): 52 t(r_3): 52	t(r_1): 52 t(r_2): 53 t(r_3): 52	t(r_1): 52 t(r_2): 52 t(r_3): 53
$k = 54$	t(r_1): 54 t(r_2): 53 t(r_3): 53	t(r_1): 53 t(r_2): 54 t(r_3): 53	t(r_1): 53 t(r_2): 53 t(r_3): 54
$k = 55$	t(r_1): 55 t(r_2): 54 t(r_3): 54	t(r_1): 54 t(r_2): 55 t(r_3): 54	t(r_1): 54 t(r_2): 54 t(r_3): 55
$k = 56$	t(r_1): 56 t(r_2): 55 t(r_3): 55	t(r_1): 55 t(r_2): 56 t(r_3): 55	t(r_1): 55 t(r_2): 55 t(r_3): 56
$k = 57$	t(r_1): 57 t(r_2): 56 t(r_3): 56	t(r_1): 56 t(r_2): 57 t(r_3): 56	t(r_1): 56 t(r_2): 56 t(r_3): 57
$k = 58$	t(r_1): 58 t(r_2): 57 t(r_3): 57	t(r_1): 57 t(r_2): 58 t(r_3): 57	t(r_1): 57 t(r_2): 57 t(r_3): 58
$k = 59$	t(r_1): 59 t(r_2): 58 t(r_3): 58	t(r_1): 58 t(r_2): 59 t(r_3): 58	t(r_1): 58 t(r_2): 58 t(r_3): 59
$k = 60$	t(r_1): 60 t(r_2): 59 t(r_3): 59	t(r_1): 59 t(r_2): 60 t(r_3): 59	t(r_1): 59 t(r_2): 59 t(r_3): 60

8 Simulation: Additional Results

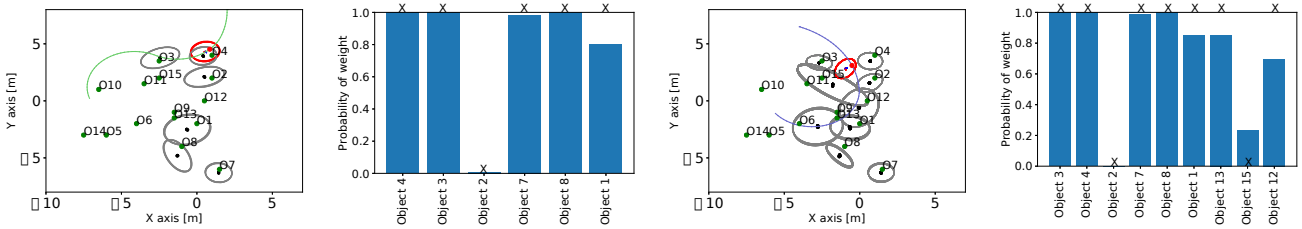
In this section we present additional results for the simulation. We show the beliefs at various stages of the path.

The results of all the graphs support the paper results, where both classification and SLAM in general are more accurate for the distributed belief. In addition, the robots inferring the distributed belief take into account objects that they didn't observe directly.



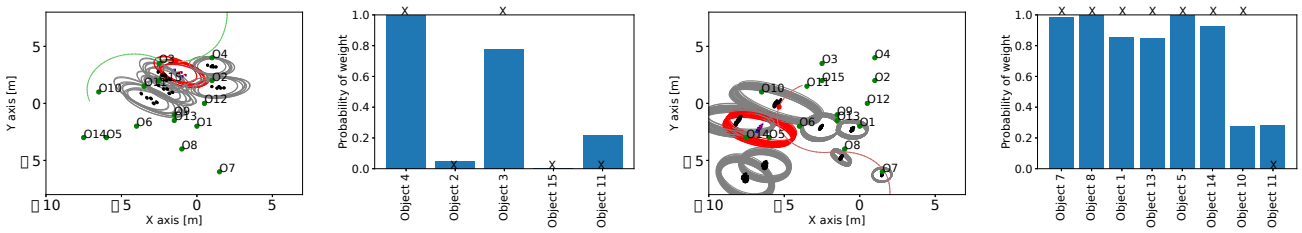
(a) Local SLAM r_2 **(b)** Local classification r_2 **(c)** Local SLAM r_3 **(d)** Local classification r_3

Figure 1: Figures for robot r_2 and r_3, local beliefs for time $k = 15$ and $k = 20$ respectively. **(a)** and **(b)** show results for r_2, **(c)** and **(d)** for r_3. **(a)** and **(c)** present SLAM results, **(b)** and **(d)** present classification results.



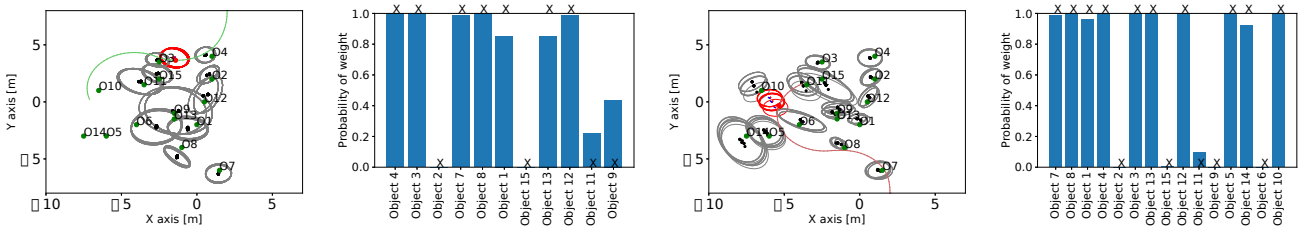
(a) Dis. SLAM r_2 **(b)** Dis. classification r_2 **(c)** Dis. SLAM r_3 **(d)** Dis. classification r_3

Figure 2: Figures for robot r_2 and r_3, distributed beliefs for time $k = 15$ and $k = 20$ respectively. **(a)** and **(b)** show results for r_2, **(c)** and **(d)** for r_3. **(a)** and **(c)** present SLAM results, **(b)** and **(d)** present classification results.



(a) Local SLAM r_2 **(b)** Local classification r_2 **(c)** Local SLAM r_1 **(d)** Local classification r_1

Figure 3: Figures for robot r_2 and r_1, local beliefs for time $k = 25$ and $k = 50$ respectively. **(a)** and **(b)** show results for r_2, **(c)** and **(d)** for r_1. **(a)** and **(c)** present SLAM results, **(b)** and **(d)** present classification results.



(a) Dis. SLAM r_2 **(b)** Dis. classification r_2 **(c)** Dis. SLAM r_1 **(d)** Dis. classification r_1

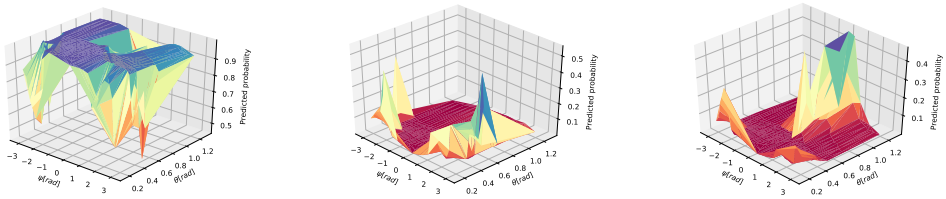
Figure 4: Figures for robot r_2 and r_1, distributed beliefs for time $k = 25$ and $k = 50$ respectively. **(a)** and **(b)** show results for r_2, **(c)** and **(d)** for r_1. **(a)** and **(c)** present SLAM results, **(b)** and **(d)** present classification results.

9 Experiment: Parameters

We consider a motion model with noise covariance $\Sigma_w = diag(0.0003, 0.0003, 0.0001)$, and geometric model with noise covariance $\Sigma_v^{geo} = diag(0.04, 0.04, 0.005)$, both corresponding to position coordinates in meters and orientation in radians. We simulated noisy odometry and geometric measurements, while using YOLO3 to create object proposals and a classifier to classify them. The communication radius in this scenario is 3 meters.

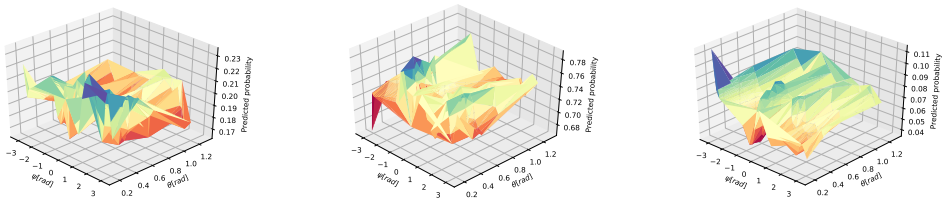
The classifier used in our experiment is the Pytorch implementation of ResNet 50, pre-trained on ImageNet dataset [2]. We trained three classifier models, one per each class. Class $c = 1$ is 'Barber Chair' and is considered our ground truth. Class $c = 2$ is 'Punching Bag' and class $c = 3$ is 'Traffic Light'. We trained the classifiers using pairs of relative pose and probability vectors; for $c = 1$, we used images of a chair used in the experiment, while for $c = 2$ and $c = 3$ we sampled measurements from Dirichlet Distribution with parameters $\alpha = [5, 15, 3]$ and $\alpha = [5, 3, 15]$ respectively. Each relative pose was parametrized by the relative yaw angle ψ, and the relative θ, with the camera being viewed from the object's frame of reference.

Fig. 5, Fig 6, and 7 present the trained expected probability values for each relative ψ and θ values, i.e. $\mathbb{P}(z^{sem}|c = i, \psi, \theta)$ for each figure with different i. Each subfigure **(a)** to **(c)** representing measurement probability of class $c = 1$ to $c = 3$ respectively.



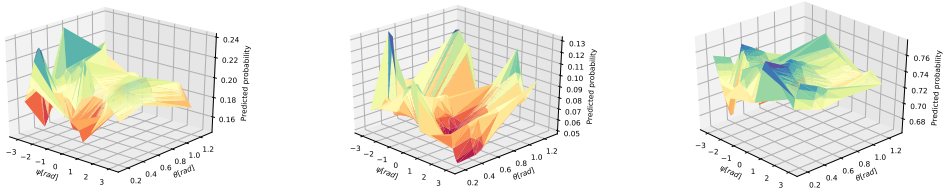
(a) $c = 1$ model, $c = 1$ prob. **(b)** $c = 1$ model, $c = 2$ prob. **(c)** $c = 1$ model, $c = 3$ prob.

Figure 5: Classifier model for $c = 1$, trained on real images: probabilities of classes 1 to 3 depending on relative yaw and pitch angles presented i **(a)** to **(c)** respectively. Higher surfaces go have bluer color.



(a) $c = 2$ model, $c = 1$ prob. **(b)** $c = 2$ model, $c = 2$ prob. **(c)** $c = 2$ model, $c = 3$ prob.

Figure 6: Classifier model for $c = 2$, trained on real images: probabilities of classes 1 to 3 depending on relative yaw and pitch angles presented i **(a)** to **(c)** respectively. Higher surfaces go have bluer color.



(a) $c = 3$ model, $c = 1$ prob. **(b)** $c = 3$ model, $c = 2$ prob. **(c)** $c = 3$ model, $c = 3$ prob.

Figure 7: Classifier model for $c = 3$, trained on real images: probabilities of classes 1 to 3 depending on relative yaw and pitch angles presented i **(a)** to **(c)** respectively. Higher surfaces go have bluer color.

10 Experiment: Table of Stack Time Stamps

In this section we present a table of stack time stamps that indicates direct and indirect communication between robots in our scenario. Recall that the maximal communication radius is 3 meters, thus all robots start to communicate between them at step $k = 6$.

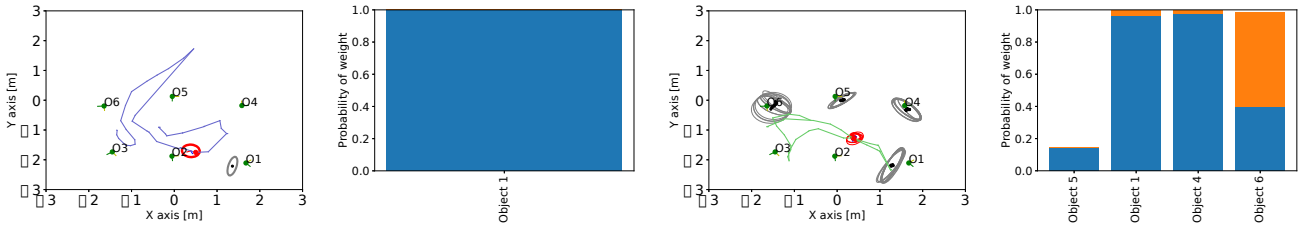
Time step	Stack of r_1	Stack of r_2	Stack of r_3
$k = 1$	t(r_1): 1 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 1
$k = 2$	t(r_1): 2 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 2
$k = 3$	t(r_1): 3 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 3 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 3
$k = 4$	t(r_1): 4 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 4 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 4
$k = 5$	t(r_1): 5 t(r_2): 0 t(r_3): 0	t(r_1): 0 t(r_2): 5 t(r_3): 0	t(r_1): 0 t(r_2): 0 t(r_3): 5
$k = 6$	t(r_1): 6 t(r_2): 5 t(r_3): 5	t(r_1): 5 t(r_2): 6 t(r_3): 5	t(r_1): 5 t(r_2): 5 t(r_3): 6
$k = 7$	t(r_1): 7 t(r_2): 6 t(r_3): 6	t(r_1): 6 t(r_2): 7 t(r_3): 6	t(r_1): 6 t(r_2): 6 t(r_3): 7
$k = 8$	t(r_1): 8 t(r_2): 7 t(r_3): 7	t(r_1): 7 t(r_2): 8 t(r_3): 7	t(r_1): 7 t(r_2): 7 t(r_3): 8
$k = 9$	t(r_1): 9 t(r_2): 8 t(r_3): 8	t(r_1): 8 t(r_2): 9 t(r_3): 8	t(r_1): 8 t(r_2): 8 t(r_3): 9
$k = 10$	t(r_1): 10 t(r_2): 9 t(r_3): 9	t(r_1): 9 t(r_2): 10 t(r_3): 9	t(r_1): 9 t(r_2): 9 t(r_3): 10
$k = 11$	t(r_1): 11 t(r_2): 10 t(r_3): 10	t(r_1): 10 t(r_2): 11 t(r_3): 10	t(r_1): 10 t(r_2): 10 t(r_3): 11
$k = 12$	t(r_1): 12 t(r_2): 11 t(r_3): 11	t(r_1): 11 t(r_2): 12 t(r_3): 11	t(r_1): 11 t(r_2): 11 t(r_3): 12
$k = 13$	t(r_1): 13 t(r_2): 12 t(r_3): 12	t(r_1): 12 t(r_2): 13 t(r_3): 12	t(r_1): 12 t(r_2): 12 t(r_3): 13
$k = 14$	t(r_1): 14 t(r_2): 13 t(r_3): 13	t(r_1): 13 t(r_2): 14 t(r_3): 13	t(r_1): 13 t(r_2): 13 t(r_3): 14
$k = 15$	t(r_1): 15 t(r_2): 14 t(r_3): 14	t(r_1): 14 t(r_2): 15 t(r_3): 14	t(r_1): 14 t(r_2): 14 t(r_3): 15

Time step	Stack of r_1	Stack of r_2	Stack of r_3
$k = 16$	t(r_1): 16 t(r_2): 15 t(r_3): 15	t(r_1): 15 t(r_2): 16 t(r_3): 15	t(r_1): 15 t(r_2): 15 t(r_3): 16
$k = 17$	t(r_1): 17 t(r_2): 16 t(r_3): 16	t(r_1): 16 t(r_2): 17 t(r_3): 16	t(r_1): 16 t(r_2): 16 t(r_3): 17
$k = 18$	t(r_1): 18 t(r_2): 17 t(r_3): 17	t(r_1): 17 t(r_2): 18 t(r_3): 17	t(r_1): 17 t(r_2): 17 t(r_3): 18
$k = 19$	t(r_1): 19 t(r_2): 18 t(r_3): 18	t(r_1): 18 t(r_2): 19 t(r_3): 18	t(r_1): 18 t(r_2): 18 t(r_3): 19
$k = 20$	t(r_1): 20 t(r_2): 19 t(r_3): 19	t(r_1): 19 t(r_2): 20 t(r_3): 19	t(r_1): 19 t(r_2): 19 t(r_3): 20
$k = 21$	t(r_1): 21 t(r_2): 19 t(r_3): 19	t(r_1): 19 t(r_2): 21 t(r_3): 20	t(r_1): 19 t(r_2): 20 t(r_3): 21
$k = 22$	t(r_1): 22 t(r_2): 19 t(r_3): 19	t(r_1): 19 t(r_2): 22 t(r_3): 21	t(r_1): 19 t(r_2): 21 t(r_3): 22
$k = 23$	t(r_1): 23 t(r_2): 19 t(r_3): 22	t(r_1): 21 t(r_2): 23 t(r_3): 21	t(r_1): 22 t(r_2): 21 t(r_3): 23
$k = 24$	t(r_1): 24 t(r_2): 19 t(r_3): 23	t(r_1): 21 t(r_2): 24 t(r_3): 21	t(r_1): 23 t(r_2): 21 t(r_3): 24
$k = 25$	t(r_1): 25 t(r_2): 23 t(r_3): 24	t(r_1): 21 t(r_2): 25 t(r_3): 24	t(r_1): 24 t(r_2): 24 t(r_3): 25
$k = 26$	t(r_1): 26 t(r_2): 24 t(r_3): 25	t(r_1): 24 t(r_2): 26 t(r_3): 25	t(r_1): 25 t(r_2): 25 t(r_3): 26
$k = 27$	t(r_1): 27 t(r_2): 25 t(r_3): 26	t(r_1): 25 t(r_2): 27 t(r_3): 26	t(r_1): 26 t(r_2): 26 t(r_3): 27
$k = 28$	t(r_1): 28 t(r_2): 26 t(r_3): 27	t(r_1): 26 t(r_2): 28 t(r_3): 27	t(r_1): 27 t(r_2): 27 t(r_3): 28
$k = 29$	t(r_1): 29 t(r_2): 27 t(r_3): 28	t(r_1): 27 t(r_2): 29 t(r_3): 28	t(r_1): 28 t(r_2): 28 t(r_3): 29
$k = 30$	t(r_1): 30 t(r_2): 29 t(r_3): 29	t(r_1): 29 t(r_2): 30 t(r_3): 29	t(r_1): 29 t(r_2): 29 t(r_3): 30

Time step	Stack of r_1	Stack of r_2	Stack of r_3
$k = 31$	t(r_1): 31 t(r_2): 30 t(r_3): 30	t(r_1): 30 t(r_2): 31 t(r_3): 30	t(r_1): 30 t(r_2): 30 t(r_3): 31
$k = 32$	t(r_1): 32 t(r_2): 31 t(r_3): 31	t(r_1): 31 t(r_2): 32 t(r_3): 31	t(r_1): 31 t(r_2): 31 t(r_3): 32
$k = 33$	t(r_1): 33 t(r_2): 32 t(r_3): 32	t(r_1): 32 t(r_2): 33 t(r_3): 32	t(r_1): 32 t(r_2): 32 t(r_3): 33
$k = 34$	t(r_1): 34 t(r_2): 33 t(r_3): 33	t(r_1): 33 t(r_2): 34 t(r_3): 33	t(r_1): 33 t(r_2): 33 t(r_3): 34
$k = 35$	t(r_1): 35 t(r_2): 34 t(r_3): 34	t(r_1): 34 t(r_2): 35 t(r_3): 34	t(r_1): 34 t(r_2): 34 t(r_3): 35
$k = 36$	t(r_1): 36 t(r_2): 35 t(r_3): 35	t(r_1): 35 t(r_2): 36 t(r_3): 35	t(r_1): 35 t(r_2): 35 t(r_3): 36
$k = 37$	t(r_1): 37 t(r_2): 36 t(r_3): 36	t(r_1): 36 t(r_2): 37 t(r_3): 36	t(r_1): 36 t(r_2): 36 t(r_3): 37
$k = 38$	t(r_1): 38 t(r_2): 37 t(r_3): 37	t(r_1): 37 t(r_2): 38 t(r_3): 37	t(r_1): 37 t(r_2): 37 t(r_3): 38
$k = 39$	t(r_1): 39 t(r_2): 38 t(r_3): 38	t(r_1): 38 t(r_2): 39 t(r_3): 38	t(r_1): 38 t(r_2): 38 t(r_3): 39
$k = 40$	t(r_1): 40 t(r_2): 39 t(r_3): 39	t(r_1): 39 t(r_2): 40 t(r_3): 39	t(r_1): 39 t(r_2): 39 t(r_3): 40

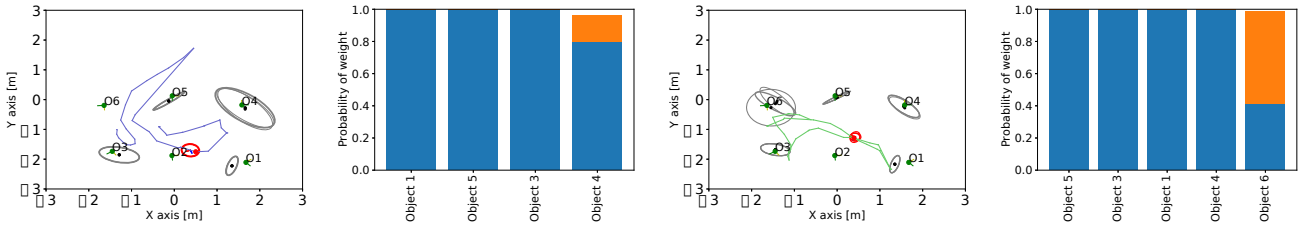
11 Experiment: Additional Results

In this section we present additional results for the experiment. We show the beliefs at various stages of the path.



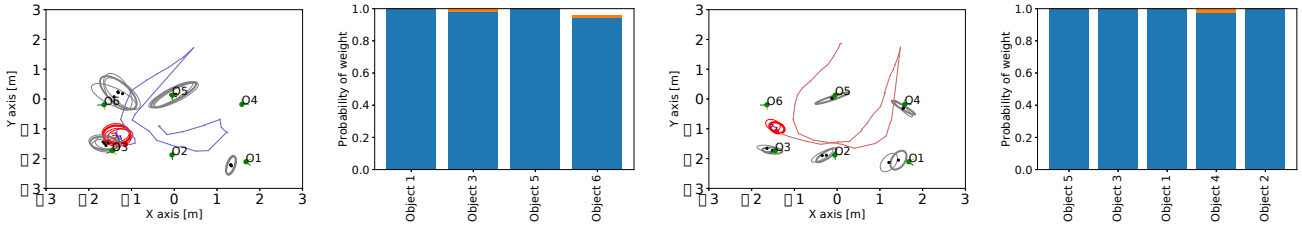
(a) Local SLAM r_3 **(b)** Local classification r_3 **(c)** Local SLAM r_2 **(d)** Local classification r_2

Figure 8: Figures for robot r_3 and r_2, local beliefs for time $k = 15$ and $k = 20$ respectively. **(a)** and **(b)** show results for r_3, **(c)** and **(d)** for r_2. **(a)** and **(c)** present SLAM results, **(b)** and **(d)** present classification results.



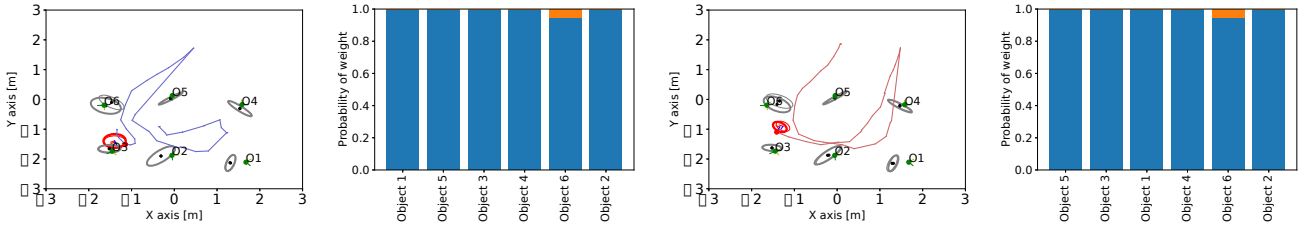
(a) Dis. SLAM r_3 **(b)** Dis. classification r_3 **(c)** Dis. SLAM r_2 **(d)** Dis. classification r_2

Figure 9: Figures for robot r_2 and r_1, distributed beliefs for time $k = 15$ and $k = 20$ respectively. **(a)** and **(b)** show results for r_3, **(c)** and **(d)** for r_2. **(a)** and **(c)** present SLAM results, **(b)** and **(d)** present classification results.



(a) Local SLAM r_1 **(b)** Local classification r_1 **(c)** Local SLAM r_3 **(d)** Local classification r_3

Figure 10: Figures for robot r_3 and r_1, local beliefs for time $k = 35$ and $k = 40$ respectively. **(a)** and **(b)** show results for r_3, **(c)** and **(d)** for r_1. **(a)** and **(c)** present SLAM results, **(b)** and **(d)** present classification results.



(a) Dis. SLAM r_3 **(b)** Dis. classification r_3 **(c)** Dis. SLAM r_1 **(d)** Dis. classification r_1

Figure 11: Figures for robot r_3 and r_1, distributed beliefs for time $k = 35$ and $k = 40$ respectively. **(a)** and **(b)** show results for r_3, **(c)** and **(d)** for r_1. **(a)** and **(c)** present SLAM results, **(b)** and **(d)** present classification results.

The results of all the graphs support the paper results as well, where both classification and SLAM in general

are more accurate for the distributed belief. In addition, the robots inferring the distributed belief take into account objects that they didn't observe directly.

References

[1] V. Tchuiev and V. Indelman. Semantic distributed multi-robot classification, localization, and mapping with a viewpoint dependent classifier model. In *IEEE Robotics and Automation Letters (RA-L)*, 2020. Submitted.

[2] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 248–255. IEEE, 2009.