

Non-parametric Bayesian State Estimator for Negative Information

Guillaume de Chambrier^{1,*}, Aude Billard¹

¹*École Polytechnique Fédérale de Lausanne (EPFL), Route Cantonale, 1015 Lausanne, Switzerland*

Correspondence*:

Guillaume de Chambrier
guillaume.dechambrier@epfl.ch

2 ABSTRACT

3 Simultaneous Localisation and Mapping (SLAM) is concerned with the development of filters to
4 accurately and efficiently infer the state parameters (position, orientation,...) of an agent and
5 aspects of its environment, commonly referred to as the map. It is necessary for that the agent
6 achieves situatedness which is a precondition planning and reasoning. In this work we consider
7 an agent who is given the task of finding a set of objects. The agent has limited perception and
8 can only sense the presences of objects if a direct contact is made. In the absence of recurrent
9 sightings or direct measurements of objects there are no correlations from the measurement
10 errors which can be exploited. This renders SLAM estimators, for which this fact is their backbone,
11 ineffective. No assumptions are taken with respect to the marginals (beliefs) of both the agent and
12 feature-based. This leads to additional restrictions on current SLAM methodologies for which the
13 parametrisation of the agents and objects joint distribution is a Multivariate Gaussian. From the
14 loose assumptions we stipulate regarding the marginals, we adopt a histogram parametrisation.
15 We introduce a SLAM algorithm, which we name Measurement Likelihood Memory Filter (MLMF),
16 in which the joint distribution is not parametrised but instead we directly apply changes from the
17 filtering step to the marginals. This is achieved by keeping track of the history of measurement
18 likelihood functions. We demonstrate that our approach gives the same filtered marginals as a
19 histogram filter and show two implementations in which both have linear space complexity with
20 one showing an exponential time complexity (although an order of magnitude smaller than the
21 histogram filter) and the other showing a linear time complexity which we call scalable-MLMF.
22 We further quantitatively demonstrate the scalability of our algorithm with 25 beliefs having up to
23 10000000 states. In an Active-SLAM setting we evaluate the impact that the size of the memory's
24 history has on the decision theoretic process in a search scenario for a one step look ahead
25 information gain planner. We report on both 1D and 2D experiments.

26 **Keywords:** Negative Information, SLAM, Bayesian State Estimator

1 INTRODUCTION

27 Estimating the location or state parameters of a mobile agent whilst simultaneously building a map of the
28 environment has been regarded as one of the most important problems to be solved for agents to achieve
29 true autonomy. It is a necessary precondition for any agent to have an estimation of the world at its disposal

30 which accurately encompasses all knowledge and related uncertainties. There has been much research
31 surrounding the field of Simultaneous Localisation And Mapping (SLAM) which branches out into a
32 wide variety of sub-fields dealing with problems from building accurate noise models of the agent sensors
33 Plagemann et al. (2007), to determining which environmental feature caused a particular measurement,
34 also known as the data association problem Montemerlo and Thrun (2003) and many more.

35 Although the amount of research might seem overwhelming at first view, all current SLAM methodologies
36 are founded on a single principle; the uncertainty of the map is correlated through the agent's measurements.
37 When an agent localises itself (by reducing position uncertainty) all previously landmarks have their
38 uncertainty reduced since the uncertainty is correlated with that of the agent's uncertainty.

39 There are three main paradigms to solving the SLAM problem. The first is EKF-SLAM (Extendend-
40 Kalman Filter) Durrant-Whyte and Bailey (2006). EKF-SLAM models the full state, being the agent's
41 parameters and environmental features, by a Multivariate Gaussian distribution. The uncertainty of each
42 individual feature is parametrised by a mean (expected position of the feature) and covariance (the level of
43 uncertainty of the position of the feature).

44 The second approach is Graph-SLAM Grisetti et al. (2010). Graph-SLAM estimates the full path of
45 the agent and considers every measurement to be a constraint on the agent's path. It is parameterised by
46 the canonical Multivariate Gaussian. At each time step a new row and column is added to the precision
47 matrix which encodes landmarks which have been observed as constraints on the robot's position. At
48 predetermined times, a nonlinear sparse optimisation is solved to minimise all the accumulated constraints
49 on the robot's path.

50 The third method is FastSLAM Montemerlo et al. (2003). FastSLAM exploits the fact that if we know
51 the agent's position with certainty all landmarks become independent. It models the distribution of the
52 agent's position by a particle filter. Each particle has its own copy of the map and updates all landmarks
53 independently which is the strength of this method. However, if many particles are required each must have
54 its own copy of the map. It is beyond the scope of this chapter to provide a detailed review of these three
55 paradigms and the reader is referred to Thrun et al. (2005), Thrun and Leonard (2008).

56 1.1 Active-SLAM & Exploration

57 Active-SLAM refers to a decision theoretic process of choosing control actions so as to actively increase
58 the convergence of the map. It is used in conjunction with exploration of an unknown environment in
59 a SLAM setting. The two steps of this process are: (i) generate a set of candidate destination positions,
60 (ii) evaluate these positions based on a utility function. The utility is a trade off between reducing the
61 uncertainty of the map or reducing the uncertainty of the agent's position.

62 Most approaches use a two-level representation of the map in an exploration setting. At the lower level
63 there is the chosen (landmark-based) SLAM filter and at the higher level a coarser representation of the
64 world. Such representations can be occupancy grids Thrun and Bü (1996) which encode either occupied
65 and free space or a topological representation Kollar and Roy (2008).

66 Early and current approaches to selecting candidate exploratory locations are based on evaluating Next-
67 best-view González-Baños and Latombe (2002) locations. Next-best-view points are sampled around *free*
68 *edges* which are at the horizon of the known map (*frontier* regions). In such a setting only target points are
69 generated, not the full trajectory. Probabilistic Road Map (PRM) Kavraki et al. (1996) based methods have
70 been used as planners to reach desired target locations, such as in Huang and Gupta (2008), where a Rapidly
71 Exploring Random Trees (RRT) is combined with FastSLAM. In Valencia et al. (2012), paths to *frontier*

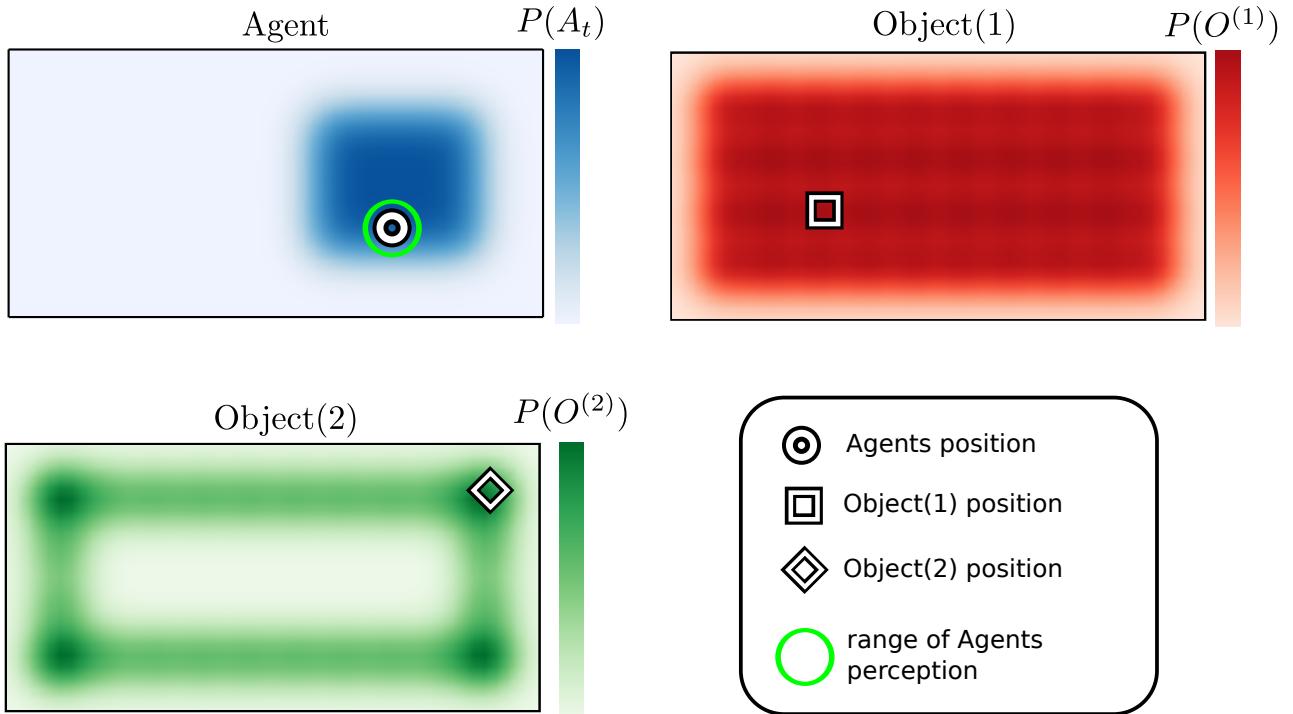


Figure 1. Table World There are three different probability density functions present on the table. The blue represents the believed location of the agent, the red and green probability distributions are associated with object 1 and 2. The white shapes in each figure represent the true location of each associated object or agent.

72 regions are computed via PRM on a occupancy grid map and at the lower level they use Pose-SLAM
73 (synonym for Graph-SLAM).

74 An alternative approach taken to generating candidate locations is the specification of high level macro
75 actions, they being either *exploratory* or *revisiting* actions as is the case in Stachniss et al. (2005). Macro
76 actions reduce the costly evaluation of actions, especially in the case of FastSLAM, which requires
77 propagating the filter forward in time so as to infer the information gain of each action.

78 The last approach is to solve the planning problem through formulating it as Partially Observable Markov
79 Decision Process (POMDP) Ross et al. (2008). However all methods take an approximation of the POMDP
80 and consider a one time step planning horizon (Lidoris, 2011, p.37).

81 There are many ways of generating actions or paths, however their utility is nearly all exclusively based
82 on the *information gain*, which is the estimated reduction of entropy a particular action or path would
83 achieve. A few utilities use f-measures such as the Kullback-Leibler divergence. Evaluation of different
84 utility metrics are presented in Carrillo et al. (2012); Carlone et al. (2010).

85 1.2 Problem Statement

86 We consider an agent searching for a set of objects in a partially-known environment, in which
87 exteroceptive feedback is extremely limited. In the case of our agent, we can think of it as having a
88 range sensor which only provides a response when in direct contact with an object. Our agent lives in a
89 *Table Top* world (see Figure 1) in which is located a set of objects. The agent's uncertainty of its location
90 and that of the objects is encoded by probability distributions $P(\cdot)$, which at initialisation are known as the
91 agent's prior beliefs.

92 As the agent explores the world, it integrates all sensing information at each time step and updates its
93 prior beliefs to posteriors (the result of the prior belief after integrating motion and sensory information).
94 All current SLAM methods are limited in that they consider only uncertainty induced by sensing inaccuracy
95 inherent in the sensor and motion models. In our setting as the sensory information is strictly haptic, we can
96 confidently assume no measurement noise. In the search task illustrated in Figure 1, the beliefs and sparse
97 measurement information available to the agent are the source of the uncertainty which is, the absence of
98 positive object measurements. This is known as **negative information** (Thrun et al., 2005, p.313) Thrun
99 (2002); Hoffman et al. (2005). Thus SLAM methodologies which use the **Gaussian error** between the
100 predicted and estimated position of features, such as in the case of EKF-SLAM and Graph-SLAM, will not
101 perform well in this setting.

102 In addition to the negative sensing information, the original beliefs depicted in Figure 1 are **non-Gaussian**
103 and **multi-modal**. We make **no assumption** regarding the form of the beliefs. This implies that the joint
104 distribution can no longer be parameterised by a Multivariate Gaussian. This is an assumption made in
105 many SLAM algorithms, notably EKF-SLAM, and allows for a closed form solution to the state estimation
106 problem. Without the Gaussian assumption no closed form solution to the filtering problem is feasible.
107 Using standard non-parametric methods (Kernel Density, Gaussian Process, Histogram,...) to represent
108 or estimate the joint distribution becomes unrealistic after a few dimensions or additional map features.
109 FastSLAM could be a potential candidate, however as it parameterises the position uncertainty of the agent
110 by a particle filter and each particle has its own copy of the map, the memory demands become quickly
111 significant. For planning purposes we would also want to have a single representation of the marginals.
112 The box below summarises the desirable attributes and assumptions for our filter.

Attributes & Assumptions

- Non-Gaussian joint distribution, no assumptions are made with respect to its form.
- Mostly negative information available (absence of positive sightings of the landmarks).
- Joint distribution volume grows exponentially with respect to the number of objects and states.
- Joint distribution volume is dense, there is high uncertainty.

113

114 1.3 The main contribution to the field of Artificial Intelligence

115 In a wide range of Artificial Intelligence (AI) applications the agent's beliefs are discrete. This non-
116 parametric representation is the most unconstraining but comes at a cost. The parameterisation of the
117 belief's joint distribution grows at the rate of a double exponential. We propose a Bayesian state estimator
118 which delivers the same filtered beliefs as a traditional filter but without explicitly parametrising the joint
119 distribution. We refer to our novel filter as the Measurement Likelihood Memory Filter (MLMF). It keeps
120 track of the history of measurement likelihood functions, referred to as the memory, which have been
121 applied on the joint distribution. The MLMF filter efficiently processes negative information. To the
122 author's knowledge there has been little research on the integration of negative information in a SLAM
123 setting. Previous work considered the case of active localisation Hoffmann et al. (2006). The incorporation
124 of negative information is useful in many contexts and in particular in Bayesian Theory of Mind, Bake

125 et al. (2011), where the reasoning process of a human is inferred from a Bayesian Network and in our own
126 work de Chambrier and Billard (2013) where we model the search behaviour of a intentionally blinded
127 human. In such a setting much negative information is present and an efficient belief filter is required. Our
128 MLMF is thus applicable to the SLAM & AI community in general and to the cognitive community which
129 models human or agent behaviours through the usage of Bayesian state estimators.

130 By using this new representation we implement a set of passive search trajectories through the state space
131 and demonstrate, for a discretised state space, that our novel filter is optimal with respect to the Bayesian
132 criteria (the successive filtered posteriors are exact and not an approximation with respect to Bayes rule).
133 We provide an analysis of the space and time complexity of our algorithm and prove that it is always more
134 efficient even when considering worst case scenarios. Lastly we consider an Active-SLAM setting and
135 evaluate how constraining the size of the number of memorised likelihood functions impacts the decision
136 making process of a greedy one-step look-ahead planner.

137 The rest of the document is structured as follows: in section 2, we overview the Bayes filter recursion
138 and apply it to a simple 1D search scenario for both a discrete and Gaussian parametrisation of the beliefs.
139 We demonstrate that discrete parametrisation gives the correct filtered beliefs but at a very high cost and
140 that the EKF-SLAM fails to provide the adequate solution. Section 3 we introduce the Measurement
141 Likelihood Memory Filter and overview its parametrisation. Section 4 we derive the computational time
142 and space complexity of the MLMF. Section 5 describes additional assumptions made with respect to the
143 MLMF to make it scalable (scalable-MLMF). In section 6 we numerically evaluate the time complexity of
144 the scalable-MLMF and check the assumption we made for it to be scalable. We investigate the filter's
145 sensitivity with respect to its parameters in an Active-SLAM setting.

2 BAYESIAN STATE SPACE ESTIMATION

146 Bayesian State Space Estimation (BSSE) focuses on incorporating observations to update a prior distribution
147 to a posterior distribution over the state space through the application of Bayes probability rules. The agent's
148 random variable, A , is associated with the uncertainty of its location in the world, the same holds for the
149 object(s') random variable(s), O . Given a sequence of actions and observations, $\{u_{1:t}, y_{0:t}\}$ (subscript $0 : t$
150 is the set from the start time $t = 0$ to the current time, $t = t$), algorithms of the BSSE family incorporate
151 this information to provide an estimate $P(A_t, O|Y_{0:t}, u_{1:t})$. This is known as the filtering problem where
152 all past information is incorporated to estimate the current state.

153 In Figure 2 we depict the general Bayesian Network (BN) of a BSSE. The BN conveys two types of
154 information, the dependence and independence relation between the random variables in the graph which
155 can be established through *d-separation* Shachter (1998). The **conditional dependence** $A \perp\!\!\!\perp O|Y$ is key to
156 all BSSE and SLAM algorithms. The strength of the dependence between the agent and object random
157 variable is governed by the measurement likelihood $P(Y_t|A_t, O)$. If the measurement likelihood does not
158 change the joint distribution, then the agent and object random variables will be independent, $A \perp\!\!\!\perp O$. If
159 they are independent, then no information acquired by the agent can be used to infer changes in the object
160 estimates.

161 We next demonstrate the behaviour of the BN joint distribution, Figure 2, for two different
162 parameterisations in the case of the absence of direct sighting of the object by the agent.

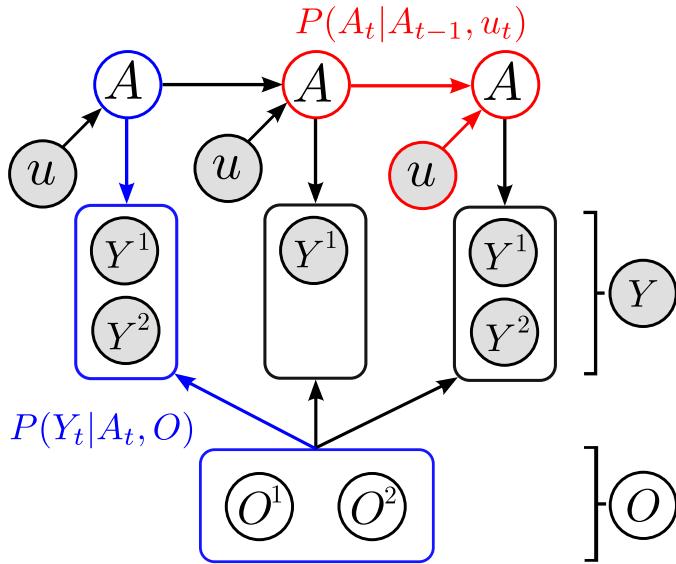


Figure 2. Directed graphical model of dependencies between the agent(A) and object(O)'s estimated location. Each object, $O^{(i)}$ is associated with one sensing random variable $Y^{(i)}$. The overall sensing random variable is $Y = [Y^{(1)}, \dots, Y^{(M-1)}]^T$, where M is the total number of agent and object random variables in the filter. For readability we have left out the time index t from A and Y . Since the objects are static, they have no temporal process associated with them thus they will never have a time subscript. The two models necessary for filtering are the motion model $P(A_t|A_{t-1}, u_t)$ (red) and measurement model $P(Y_t|A_t, O)$ (blue).

163 2.1 EKF-SLAM

164 In EKF-SLAM the joint density $p(A_t, O|Y_{0:t}, u_{1:t}) = g(x; \mu_t, \Sigma_t)$ is parametrised by a single Gaussian
 165 function g with mean, $\mu_t = [\mu_{A_t}, \mu_{O^{(1)}}, \dots, \mu_{O^{(M-1)}}]^T \in \mathbb{R}^{3+2\cdot(M-1)}$ where the random variables are in
 166 \mathbb{R}^2 , and covariance, Σ_t . The mean value of the agent $\mu_a = [x, y, \phi]^T \in \mathbb{R}^3$ and those of the objects are
 167 $\mu_{O^{(i)}} = [x, y]^T \in \mathbb{R}^2$.

$$\Sigma_t = \begin{bmatrix} \Sigma_a & \Sigma_{ao} \\ \Sigma_{oa} & \Sigma_o \end{bmatrix} \in \mathbb{R}^{(3+2\cdot(M-1)) \times (3+2\cdot(M-1))} \quad (1)$$

The j 'th object measurement is described by range and bearing $Y_t^{(j)} = [r, \phi]$ in the frame of reference of the agent. EKF-SLAM assumes that the measurement is corrupted by Gaussian noise, $\epsilon \sim \mathcal{N}(0, R)$, resulting in the likelihood function:

$$p(Y_t|A_t, O_t) = \frac{1}{|2\pi R|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} (Y_t - \hat{Y}_t)^T R^{-1} (Y_t - \hat{Y}_t) \right) \quad (2)$$

$$\hat{Y}_t = \exp \left(-\frac{1}{2\sigma^2} \|A_t - O\|^2 \right) \quad (3)$$

168 where the covariance, R , encompasses the uncertainty in the measurement and Equation 3 is the
 169 measurement function. The elements of the covariance matrix capture the measurement error between the
 170 true Y and expected \hat{Y} range and bearing of the object. As the joint distribution is parametrised by a single

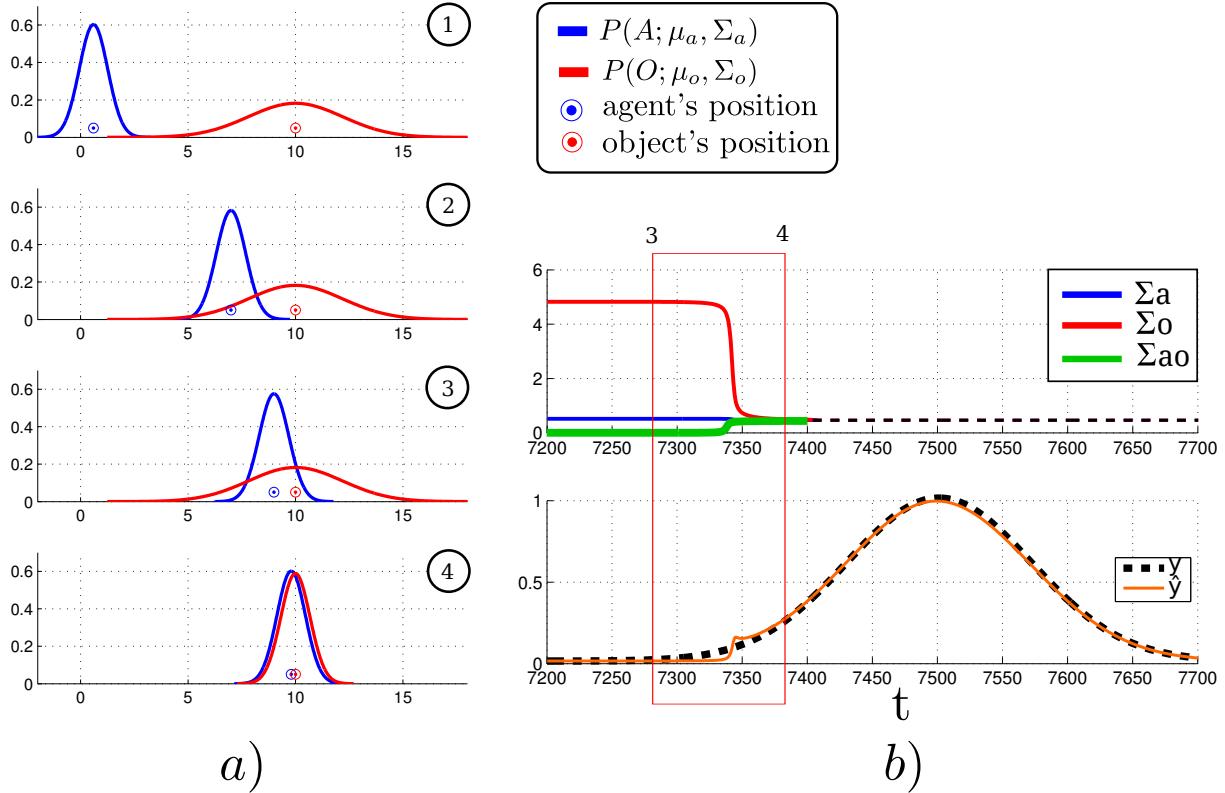


Figure 3. **a)** EKF-SLAM time slice evolutions of the pdfs. The temporal ordering is given by the numbers in the top right corner of each plot. The blue pdf represents the agent's believed location and the circle is the agent's true location. The same holds for the red distribution which represents the agent's belief of the location of an object. **b)** Evolution of the covariance components of Σ over time and true Y_t and expected measurements, \hat{Y}_t . Σ_a and Σ_o are the variances of the agent and object positions and Σ_{ao} is the cross-covariance term.

171 Multivariate Gaussian, a closed form solution to the filtering Equations exists, called the Kalman Filter
 172 Durrant-Whyte and Bailey (2006).

173 The error between the true and expected measurement $e = (Y_t - \hat{Y}_t)$ is an important part of the application
 174 of EKF-SLAM. In our scenario the agent can only perceive the objects once he enters in direct contact
 175 with them. This means that the variance of the observation Y_t will always be equal to \hat{Y} until a contact
 176 occurs. To illustrate the problems which this gives rise to, we give an illustration of a 1D search. Figure 3
 177 shows the resulting updates of the beliefs for 4 chosen time segments.

178 As expected we do not get the desired behaviour, that the beliefs start updating as soon as they are
 179 overlapping, see 2nd-3rd temporal snapshot in the Figure. Even when most of the belief mass of the
 180 agent's location pdf overlaps that of the object pdf, no belief update occurs. The multivariate Gaussian
 181 parameterisation only guarantees a dependency between the agent and object random variables when there
 182 is a positive sighting of the landmarks. This can be seen in Figure 3 (b), where the component Σ_{ao} is 0
 183 most of the time which implies that $A \perp\!\!\!\perp O|Y$ which is undesirable.

184 2.2 Histogram-SLAM

185 In Histogram-SLAM, the joint distribution is discretized and each bin has a parameter,
 186 $P(A_t = i, O = j | Y_{0:t}, u_{1:t}; \theta) = \theta^{(ij)}$, which sums to one, $\sum_{ij} \theta^{(ij)} = 1$. For shorthand notation we
 187 will write $P(A_t, O | Y_{0:t}, u_{1:t}; \theta)$ instead of $P(A_t = i, O = j | Y_{0:t}, u_{1:t}; \theta)$. The probability distribution of the

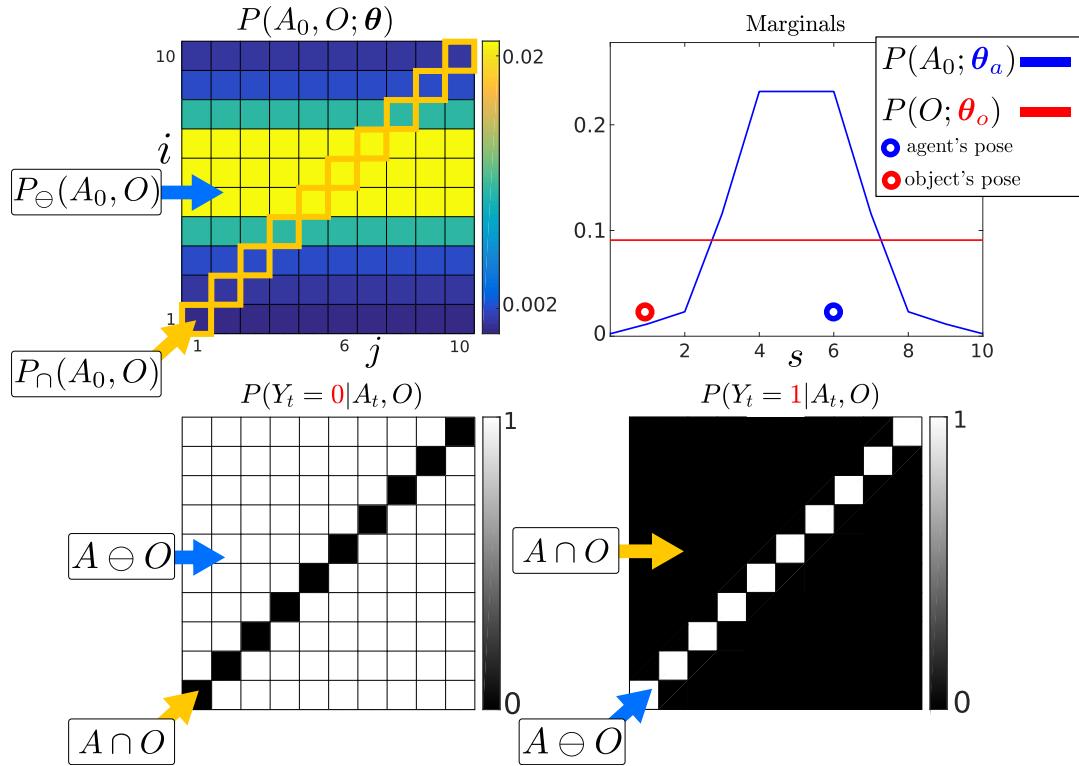


Figure 4. **Top:** *Left:* Initialisation of the agent and object joint distribution. *Right:* Marginals of the agent and object parameterised by θ_a and θ_o , giving the probability of their location. The marginal of each random variable is obtained from Equation 4. The probability of the agent and object being in state $s = 6$ is given by summing the blue and red highlighted parameters in the joint distribution. **Bottom:** 1D world Likelihood $P(Y_t|A_t, O)$, the white regions $A \cap O$ will leave the joint distribution unchanged whilst the black regions will evaluate the joint distribution to zero. *Left:* No contact detected with the object, the current measurement is $Y_t = 0$, both the agent and object cannot be in the same state. *Right:* The agent entered into contact with the object and received a haptic feedback $Y_t = 1$. The agent receives only two measurement possibilities, contact or no contact.

188 agent's position is given by marginalising the object random variable:

$$P(A_t|Y_{0:t}, u_{1:t}; \theta_a) = \sum_{j=1}^{|O|} P(A_t, O = j|Y_{0:t}, u_{1:t}; \theta) \quad (4)$$

189 The converse holds true for the object's marginal, that is Equation 4 summation would be over the agents
190 variable . Figure 4 (*Top*) illustrates the joint distribution of both the agent and the object random variable.
191 The 1D world of the agent and object is discretised to 10 states, producing a joint distribution with 100
192 parameters! For a state space of N bins, $i = 1\dots N$, and there is a total of M random variables (one
193 agent and $M - 1$ objects) and the joint distribution has N^M parameters. This exponential increase renders
194 Histogram-SLAM intractable with this parameterisation.

195 In the tasks we consider, an observation occurs only if the agent enters in contact with the object, which
196 implies that both occupy the same discrete state. The likelihood function $P(Y_t|A_t, O)$ is:

$$P(Y_t = 1|A_t, O) = \begin{cases} 1 & \text{if } A_t = O \\ 0 & \text{if } A_t \neq O \end{cases} \quad (5)$$

197 Figure 4 (*Bottom*), illustrates the likelihood of Equation 5 in the case when a no contact measurement
 198 $Y_t = 0$ is present in a 1D world. Both likelihoods are sparse in the sense that there is only a small region
 199 which gets affected by the likelihood. When there is no measurement (*Left subfigure*) all the parameters of
 200 the joint distribution which are in the black regions become zero, which we refer to as the **dependent states**
 201 $A \cap O$ of the joint distribution. The white states are the **independent states** $A \ominus O$, they are not changed by
 202 the likelihood function and the values of the joint distribution in those states, $P_{\cap}(A_t, O|Y_{0:t}, u_{1:t})$, will be
 203 unchanged $P_{\ominus}(A_t, O|Y_{0:t}, u_{1:t}) \propto P_{\ominus}(A_t, O|Y_{0:t-1}, u_{1:t})$. When the object is detected (*Right subfigure*)
 204 the likelihood constrains all non-zero values of the joint distribution to be in states $i = j$, which in the case
 205 of a 2-dimensional joint distribution is a line. The **sparsity** of the likelihood function will be key to the
 206 development of the MLMF filter. Two models are needed to perform the recursion, namely the motion
 207 model $P(A_t|A_{t-1}, u_t)$ and the measurement model $P(Y_t|A_t, O)$, which we already detailed. Both models
 208 applied consecutively to the initial joint distribution results in a posterior distribution. Both Equation 7-8
 209 are part of the histogram Bayesian filter update:

Histogram Bayesian recursion

initialisation

$$P(A_0, O; \boldsymbol{\theta}) = P(A_0; \boldsymbol{\theta}_a) P(O; \boldsymbol{\theta}_o) = \boldsymbol{\theta}_a \times \boldsymbol{\theta}_o \quad (6)$$

motion

$$P(A_t, O|Y_{0:t-1}, u_{1:t}) = \sum_{A_{t-1}} P(A_t|A_{t-1}, u_t) P(A_{t-1}, O_t|Y_{0:t-1}, u_{1:t-1}) \quad (7)$$

measurement

$$P(A_t, O|Y_{0:t}, u_{1:t}) = \frac{P(Y_t|A_t, O) P(A_t, O|Y_{0:t-1}, u_{1:t})}{P(Y_t|Y_{0:t-1}, u_{1:t})} \quad (8)$$

210

211 Figure 5 illustrates the evolution of the joint distribution in a 1D example. The agent and object's true
 212 positions (unobservable) are in state 6 and 1. The agent moves three steps towards state 10. At each time
 213 step, as the agent fails to sense the object, the likelihood function $P(Y_t = 0|A_t, O)$ (Figure 4, *Bottom right*)
 214 is applied. As the agent moves towards the right, the motion model shifts the joint distribution towards
 215 state 10 along the agent's dimension, i (note that state 1 and 10 are wrapped).

216 As the agent moves to the right more joint distribution parameters become zero. The re-normalisation by
 217 the **evidence** ($P(Y_t|Y_{0:t-1}, u_{1:t})$, denominator of Equation 8), which increases the value of the remaining
 218 parameters, is equal to the sum of the probability mass which was set to zero by the likelihood function,
 219 thus the values of the parameters of the joint distribution which fall on the pink line in Figure 5 (green line
 220 also, but only for first time slice) become zero and their values are redistributed to the remaining non-zero
 221 parameters. This is an **important aspect** which will be present in MLMF.

222 The **inconvenience** with Histogram-SLAM is that its time and space complexity is exponential as the
 223 joint distribution is discretised and parametrised by $\boldsymbol{\theta}^{(ij)}$. Instead we propose a new filter, MLMF, which
 224 we formally introduce in the next section. This filter achieves the same result as the Histogram filter but
 225 without having to parameterise the joint distribution, thus avoiding the exponential growth cost.

226 The **key idea** behind the mechanism of the MLMF filter is to evaluate only the joint distribution
 227 $P_{\cap}(A_t, O|Y_{0:t}, u_{1:t})$ in dependent states and updates directly to the marginals without parameterising the

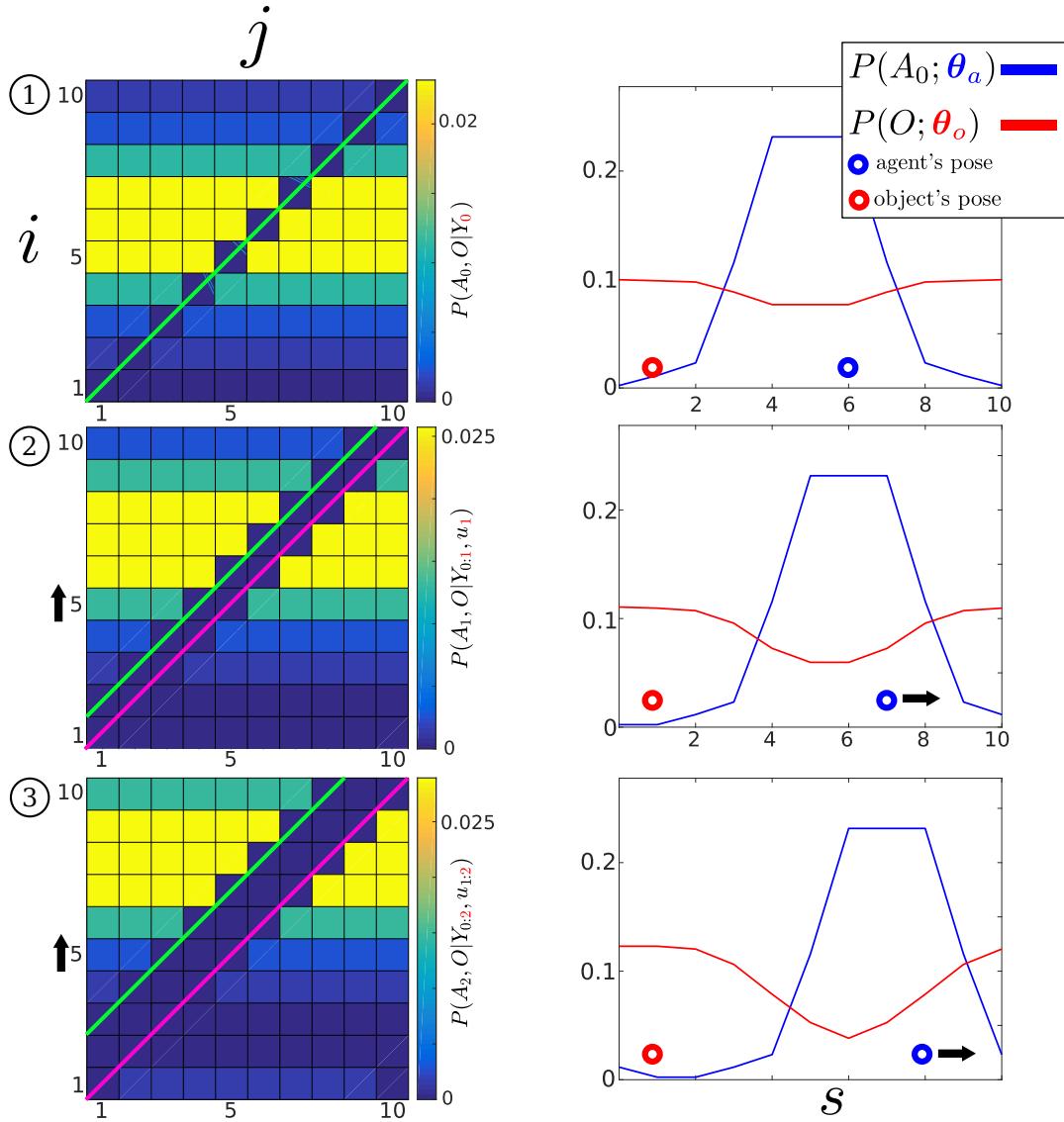


Figure 5. Histogram-SLAM, 4 time steps. **1** Application of likelihood $P(Y_0 = 0 | A_0, O)$ and the agent remains stationary, all states along the green line become zero. **2** The agent moves to the right $u_1 = 1$, the motion $P(A_1|A_0, u_1)$, and likelihood models are applied consecutively. The right motion results in a shift (black arrow on the left) in the joint probability distribution towards the state $i = 10$. All parameters on the pink line are zero. **3** Same as two. At each time step a new likelihood function (pink line) is applied to the joint distribution.

values of the joint distribution. The MLMF filter parametrises **explicitly** the marginals $P(A_t|Y_{0:t}, u_{1:t}; \theta_a)$, $P(O|Y_{0:t}, u_{1:t}; \theta_o)$. This contrasts the Histogram filter where the marginals are derived from the joint distribution by marginalisation over the entire joint state space.

3 MEASUREMENT LIKELIHOOD MEMORY FILTER

The likelihood function $P(Y_t|A_t, O)$ is the cause of the dependence between the random variables. If both the agent and object were completely independent, no additional parameters would be required to represent the joint distribution other than the marginals $P(A_t|Y_{0:t}, u_{1:t})$ and $P(O|Y_{0:t})$ giving a total of NM parameters (where N is the number of states and M is the number of random variables). At the other

235 extreme if every single point in the domain of the random variables was dependent this would require the
 236 totality of N^M parameters as previously stated in the case of the histogram filter. We propose a method in
 237 which we **do not parameterise the values of the joint distribution** but rather only compute its impact
 238 on the marginals.

239 **3.1 MLMF parametrisation**

240 MLMF keeps a **function parameterisation** of the joint distribution instead of a **value parameterisation**
 241 as it is the case for Histogram-SLAM. At initialisation the joint distribution is represented by the product
 242 of marginals, Equation 9, which would result in the joint distribution illustrated in Figure 4, if it were to
 243 be evaluated at all states (i, j) as it is done for Histogram-SLAM, Equation 6. MLMF will only evaluate
 244 this product, when necessary, at specific states. At each time step the motion and measurement update are
 245 applied, Equation 10-11. An important distinction is that these updates are performed on the **unnormalised**
 246 joint distribution, which is not the case in Histogram-SLAM where the updates are done on the conditional,
 247 Equation 7-8. After applying multiple motion and measurement updates the resulting joint distribution is
 248 given by Equation 12, see Appendix 8.3 for a step-by-step derivation.

MLMF Bayesian filter

joint marginals (initial)

$$P(A_0, O) = P(A_0; \theta_a^*) P(O; \theta_o^*) \quad (9)$$

motion

$$P(A_t, O, Y_{0:t-1}|u_{1:t}) = \sum_{A_{t-1}} P(A_t|A_{t-1}, u_t) P(A_{t-1}, O, Y_{0:t-1}|u_{1:t-1}) \quad (10)$$

measurement

$$P(A_t, O, Y_{0:t}|u_{1:t}) = P(Y_t|A_t, O) P(A_t, O, Y_{0:t-1}|u_{1:t-1}) \quad (11)$$

joint

$$P(A_t, O|Y_{0:t}, u_{1:t}; \theta_o^*, \theta_a^*, \Psi_{0:t}, \alpha_{0:t}) = \frac{P(O; \theta_o^*) P(A_t|u_{1:t}; \theta_a^*) P(Y_{0:t}|A_t, O, u_{1:t}; \Psi_{0:t})}{P(Y_{0:t}|u_{1:t}; \alpha_{0:t})} \quad (12)$$

filtered marginal

$$P(A_t|Y_{0:t}; \theta_a) = P(A_t|Y_{0:t-1}; \theta_a) - \left(P_{\cap}(A_t|Y_{0:t-1}) - P_{\cap}(A_t|Y_{0:t}) \right) \quad (13)$$

$$P(O|Y_{0:t}; \theta_o) = P(O|Y_{0:t-1}; \theta_o) - \left(P_{\cap}(O|Y_{0:t-1}) - P_{\cap}(O|Y_{0:t}) \right) \quad (14)$$

249

250 The MLMF filter, Equation 12, is parameterised by the agent and object **joint marginals**
 251 $P(A_t|u_{1:t}; \theta_a^*)$, $P(O; \theta_o^*)$, the **filtered marginals** $P(A_t|Y_{0:t}, u_{1:t}; \theta_a)$ ($u_{1:t}$ not shown for easing the
 252 notation), $P(O|Y_{0:t}; \theta_o)$, the evidence $P(Y_{0:t}|u_{1:t}; \alpha_{0:t})$ and the history of likelihood functions,
 253 $P(Y_{0:t}|A_t, O, u_{1:t}; \Psi_{0:t})$ Equation 15, which is the product of all the likelihood functions since $t = 0$ until
 254 t and we will refer to it as the **memory likelihood function**:

$$P(Y_{0:t}|A_t, O, u_{1:t}; \Psi_{0:t}) := \prod_{i=0}^t P(Y_i|A_t, O, u_{i+1:t}; l_i) \quad (15)$$

$$P(Y_i = 0|A_t, O, u_{i+1:t}; l_i) := \begin{cases} 0 & \text{if } A_t + l_i = O \\ 1 & \text{else} \end{cases} \quad (16)$$

255

$$l_i := \sum_{j=i+1}^t u_j \quad (17)$$

256 The memory likelihood function's parameters $\Psi_{0:t} = \{(Y_i, l_i)\}_{i=0:t}$ consist of a set of measurements
 257 $Y_{0:t}$ and offsets $l_{0:t}$ depicted in greed. The measurements $Y_i \in \{0, 1\}$ are always binary, whilst the offsets l_i ,
 258 actions u_t , agent A_t and object O variables' size are equal to the dimension of the state space. The subscript
 259 i of an offset l_i indicates which likelihood function it belongs to. The offset of a likelihood function is
 260 given by the summation of all the applied actions from the time the likelihood was added until the current
 261 time t , Equation 17, which can be computed recursively. The motion update, Equation 10, when applied to
 262 the joint distribution results in the initial marginal $P(A_0; \theta_a^*)$ and the likelihood functions being moved
 263 along the agent's axis. In Algorithm 1, we detail how an action u_t and measurement Y_t , result in the update
 264 of the memory likelihood's parameters from $\Psi_{0:t-t}$ to $\Psi_{0:t}$; this is an implementation of Equations 10-11.

Algorithm 1: Memory Likelihood update

input : $\Psi_{0:t-1}, Y_t, u_t$
output : $\Psi_{0:t}$

265 **motion update** $\bar{\Psi}_{0:t} \leftarrow \Psi_{0:t-1}$
 1 **for** $l_i \in \Psi_{0:t-1}$ **do**
 2 $l_i = l_i + u_t$

measurement update
 3 $\Psi_{0:t} \leftarrow \{\bar{\Psi}_{0:t}, (Y_t, l_t := 0)\}$

266 Figure 6 illustrates the evolution of the non-normalised MLMF joint distribution, Equation ???. For
 267 ease of notation we will omit at times the parameters of the probability functions. Both $P(A_0; \theta_a^*)$ and
 268 $P(O; \theta_o^*)$ were initialised as for the Histogram-SLAM example in Figure 5 on page 10. Two actions
 269 $u_{1:2} = 1$ are applied and three measurements $Y_{0:2} = 0$ received which are then integrated into the filter.
 270 Since initialisation of the joint distribution at $t = 0$ until $t = 2$ the object's marginal $P(O; \theta_o^*)$ remains
 271 unchanged and the agent's marginal $P(A_2|u_{1:2}; \theta_a^*)$ is updated by the two actions according to the motion
 272 update, see Figure 6 *Top-right*. The product of these two marginals (terms of Equation ?? before the
 273 memory likelihood product) results in the joint probability distribution $P(A_2, O|u_{1:2}; \theta_a^*, \theta_o^*)$ illustrated in
 274 Figure 6 *Middle-right*.

275 In the left column of Figure 6 we illustrate how the memory likelihood term, Equation 15, is updated
 276 according to Algorithm 1. In the *Top-left*, the first likelihood function $P(Y_0|A_2, O, u_{1:2}; l_0)$ is illustrated.
 277 As two actions have been applied, Algorithm 1 is applied twice which results in a $l_0 = 2$ parameter for the
 278 first likelihood function. In the figure we highlighted the likelihood in light-green to indicate that it was

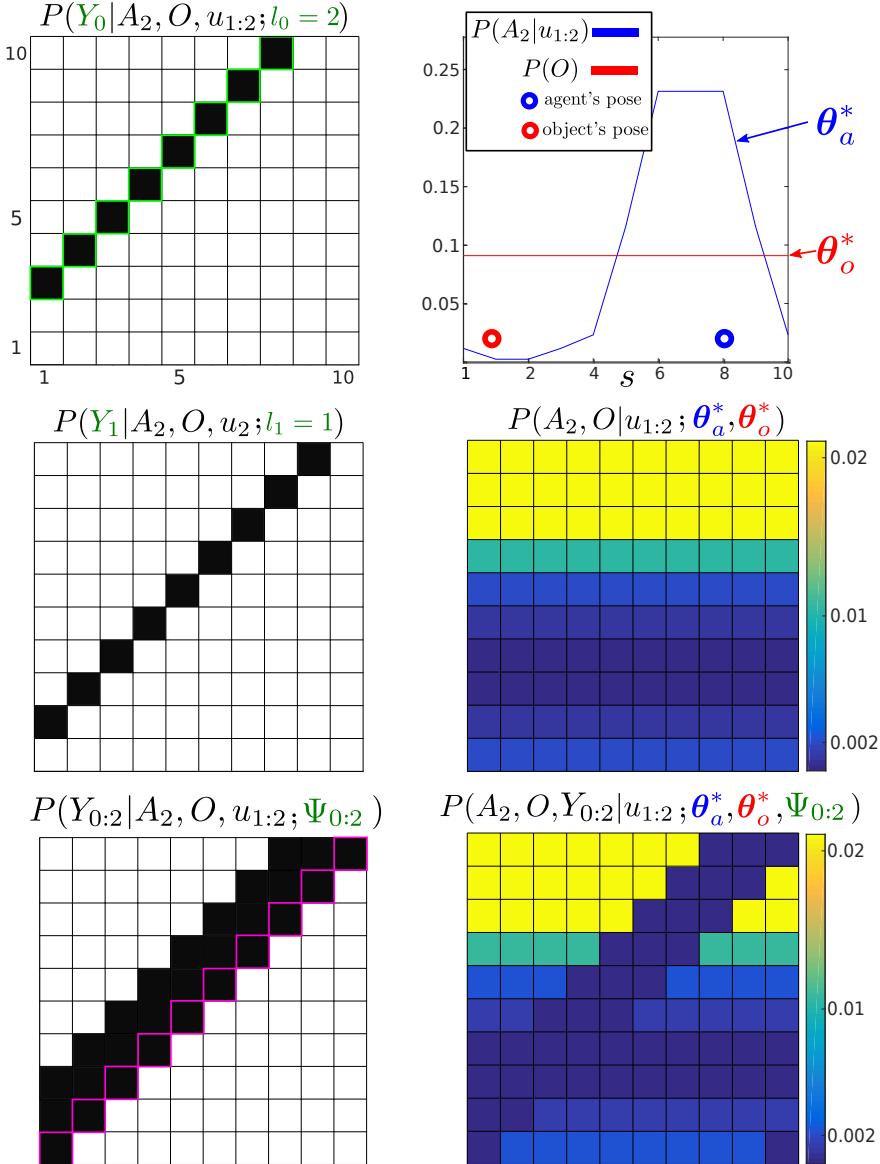


Figure 6. Unnormalised MLMF joint distribution, numerator of Equation 12, at time $t = 2$. Three measurements (all $Y = 0$) and two actions (both $u = 1$) have been integrated into the joint distribution. *Left column:* The first plot illustrates the likelihood of the first measurement Y_0 . We highlight the contour in light-green to indicate that it was the first applied likelihood function (see the correspondence with Figure 5). The first likelihood function has been moved by the 2 actions, the second likelihood function has been moved by one action (the last one, $u_2 = 1$) and the third likelihood has had no action applied to it yet. The last applied likelihood function is highlighted in pink and the product of all the likelihoods since $t = 0$ until $t = 3$ is depicted at the bottom of the figure which is $P(Y_{0:2}|A_2, O, u_{1:2})$. *Right column:* the top figure illustrates the original marginal of the object $P(O)$, which remains unchanged, and the agent's marginal $P(A_2|u_{1:2})$ which has moved in accordance to the motion update function. Their product would result in the joint distribution $P(A_2, O|u_{1:2})$ illustrated in the middle figure if evaluated at each state (i, j) . The bottom figure is the result of multiplying $P(A_2, O|u_{1:2})$ with $P(Y_{0:2}|A_2, O, u_{1:2})$, giving the numerator of Equation 12.

279 the first added to the memory term making it convenient to compare to Figure 5 on page 10. As for the
 280 second likelihood function $P(Y_1|A_2, O, u_2; l_1)$ only one action has been applied and the third likelihood
 281 function $P(Y_2|A_2, O; l_2 = 0)$ has not yet been updated by the next action. The parameters of the memory

282 likelihood function, Equation 15, are: $\Psi_{0:2} = \{(0, 2)_{i=0}, (0, 1)_{i=1}, (0, 0)_{i=2}\}$ and the evaluation of memory
 283 likelihood is depicted in the *Bottom-left* of Figure 6.

284 The reader may have noticed that the amplitude of the values of the filtered joint distribution illustrated in
 285 Figure 6 have changed when compared with Figure 5, but not the structure. This is because we have not
 286 re-normalised the joint distribution by the evidence $P(Y_{0:t}|u_{1:t}; \alpha_{0:t})$. We will show in the next section how
 287 we can **recursively** compute the evidence without having to integrate the whole joint distribution which
 288 would be expensive.

289 Our goal is to be able to compute the marginals $P(A_t|Y_{0:t}, u_{1:t}; \theta_a)$, $P(O|Y_{0:t}; \theta_o)$ of the agent and object
 290 random variables and evidence $P(Y_{0:t}|u_{1:t}; \alpha_{0:t})$ **without** having to perform an **expensive marginalisation**
 291 over the entire space of the joint distribution as was the case for Histogram-SLAM. The next section
 292 describes how to efficiently compute the evidence and the marginals. For ease of notation we will not
 293 always show the conditioned actions $u_{1:t}$, so $P(A_t, O|Y_{0:t}, u_{1:t})$ will be $P(A_t, O|Y_{0:t})$.

294 3.2 Evidence and marginals

295 In order to compute efficiently the marginal likelihood (also known as evidence) $P(Y_{0:t}|u_{1:t}; \alpha_{0:t})$ and the
 296 filtered marginals $P(A_t|Y_{0:t}, u_{1:t}; \theta_a)$, $P(O|Y_{0:t}; \theta_o)$ we take advantage of the fact that only a very small
 297 area in the joint distribution space will be affected by the measurement likelihood function at each time
 298 step. Without lost of generality the likelihood function will only make a difference to dependent $A \cap O$
 299 states in the joint distribution, states where the likelihood function is less than one. The states inside $A \ominus O$
 300 will not be affected, where the likelihood function is equal to one.

$$P(A_t, O|Y_{0:t}) = P_{\cap}(A_t, O|Y_{0:t}) + P_{\ominus}(A_t, O|Y_{0:t}) \quad (18)$$

301 This formulation will lead to large computational gain as the independent term is not influenced by the
 302 measurement function: $P_{\ominus}(A_t, O, Y_{0:t}) = P_{\ominus}(A_t, O, Y_{0:t-1})$ and $P_{\ominus}(A_t, O|Y_{0:t}) \propto P_{\ominus}(A_t, O|Y_{0:t-1})$.

303 3.2.1 Evidence

304 The evidence of the measurement $P(Y_{0:t}|u_{1:t}; \alpha_{0:t})$ is the normalisation coefficient of the joint distribution
 305 Equation 12. It is the amount of probability mass re-normalised to the other parameters as a result of the
 306 consecutive application of the likelihood function. At time step t , the normalising factor to be added to
 307 the evidence is the difference between the probability mass located inside $A \cap O$ before and after the
 308 application of the measurement function $P(Y_t|A_t, O)$, see Equation 19-20 (see Appendix 8.4 for the full
 309 derivation).

$$\alpha_t = \sum_{A_t} \sum_O \left(P(Y_t|A_t, O) - 1 \right) P_{\cap}(A_t, O, Y_{0:t-1}|u_{1:t}) \quad (19)$$

$$P(Y_{0:t}|u_{1:t}; \alpha_{0:t}) = 1 + \underbrace{\alpha_{0:t-1}}_{\alpha_{0:t}} + \alpha_t \quad (20)$$

310 The advantage of Equation 19 is that the summation is only over the states which are in the dependent area
 311 \cap of the joint distribution. This is generally always much smaller than the full space itself. Until an object is
 312 sensed, the likelihood will always be zero $P(Y_t|A_t, O) = 0$ and α_t will correspond to the probability mass
 313 which falls within the region of the joint distribution in which the likelihood function is zero. The point of

314 interest is that as we perform the filtering process we will never re-normalise the whole joint distribution,
 315 but only keep track of how much it should have been normalised. To this end the marginals $P(A_t|u_{1:t}; \theta_a^*)$
 316 and $P(O; \theta_o^*)$ are never re-normalised but are used at each step to compute how much of the probability
 317 mass α_t should go to the normalisation factor $P(Y_{0:t}|u_{1:t}; \alpha_{0:t})$. The normalisation factor in question will
 318 never be negative, as the joint distribution sums to one and each α_t represents some of the mass removed
 319 from the joint distribution. Since we keep track of the history of applied measurement likelihood functions
 320 the same amount of probability mass is never removed twice from the joint distribution.

321 3.2.2 Marginals

322 There are two different sets of marginals used in the MLMF filter. The first set are the **joint marginals** of
 323 the joint distribution, Equation 12 parameterised by θ_a^* and θ_o^* . The second set of marginals are the **filtered**
 324 **marginals** which are updated by evaluating the joint distribution in dependent states and are parameterised
 325 by θ_a and θ_o . At initialisation before the the first action or observation is made the parameters of the
 326 filtered marginal are set equal to those of the joint distribution.

327 In Histogram-SLAM both the agent and object marginals are obtained, at each time step, by marginalising
 328 the joint distribution. This requires storing and summing over all the parameters of the joint distribution
 329 which is expensive. Instead the MLMF takes advantage of the sparsity of the likelihood function which
 330 results in only the dependent elements of the marginal being affected, Equation 21 (see Appendix 8.5 for
 331 the full derivation of Equation 21).

$$P(A_t|Y_{0:t}, u_{1:t}; \theta_a) = P(A_t|Y_{0:t-1}, u_{1:t}; \theta_a) - \left(P_{\cap}(A_t|Y_{0:t-1}, u_{1:t}) - \mathbf{P}_{\cap}(\mathbf{A}_t|\mathbf{Y}_{0:t}, \mathbf{u}_{1:t}) \right) \quad (21)$$

$$\begin{aligned} \mathbf{P}_{\cap}(\mathbf{A}_t|\mathbf{Y}_{0:t}, \mathbf{u}_{1:t}; \theta_a^*, \theta_o^*, \Psi_{0:t}, \alpha_{0:t}) &= \sum_O \mathbf{P}_{\cap}(\mathbf{A}_t, \mathbf{O}|\mathbf{Y}_{0:t}, \mathbf{u}_{1:t}; \theta_a^*, \theta_o^*, \Psi_{0:t}, \alpha_{0:t}) \\ &= \frac{\sum_O P_{\cap}(O; \theta_o^*) P_{\cap}(A_t|u_{1:t}; \theta_a^*) P(Y_{0:t}|A_t, O, u_{1:t}; \Psi_{0:t})}{P(Y_{0:t}|u_{1:t}; \alpha_{0:t})} \end{aligned} \quad (22)$$

332 Equation 21 is recursive, $P(A_t|Y_{0:t}, u_{1:t}; \theta_a)$ is computed in terms of $P(A_t|Y_{0:t-1}, u_{1:t}; \theta_a)$. Figure 7
 333 illustrates a measurement update of the MLMF. The illustrated marginals (*Bottom row*) are the **filtered**
 334 **marginals** $P(A_t|Y_{0:t}, u_{1:t}; \theta_a)$, $P(O|Y_{0:t}; \theta_o)$. The shape of the **joint marginals** $P(A_t|u_{1:t}; \theta_a^*)$, $P(O; \theta_o^*)$
 335 (not shown here, see Figure 6) remain unchanged by measurements during the filtering process. They
 336 are the parameters of the joint distribution used to update the filtered marginals. Table 1 summarises the
 337 functions and parameters of the MLMF for two random variables, an agent and object.

338 We evaluate this formulation of the joint distribution with the standard histogram filter in the case of the
 339 1D filtering scenario illustrated in Figure 5 on page 10 and we find them to be identical. Having respected
 340 the formulation of Bayes rule, we assert that Algorithm 2 is a Bayesian Optimal Filter¹. Next we evaluate
 341 both space and time complexity of the MLMF filter.

¹ An optimal Bayesian solution is an exact solution to the recursive problem of calculating the exact posterior density Arulampalam et al. (2002)

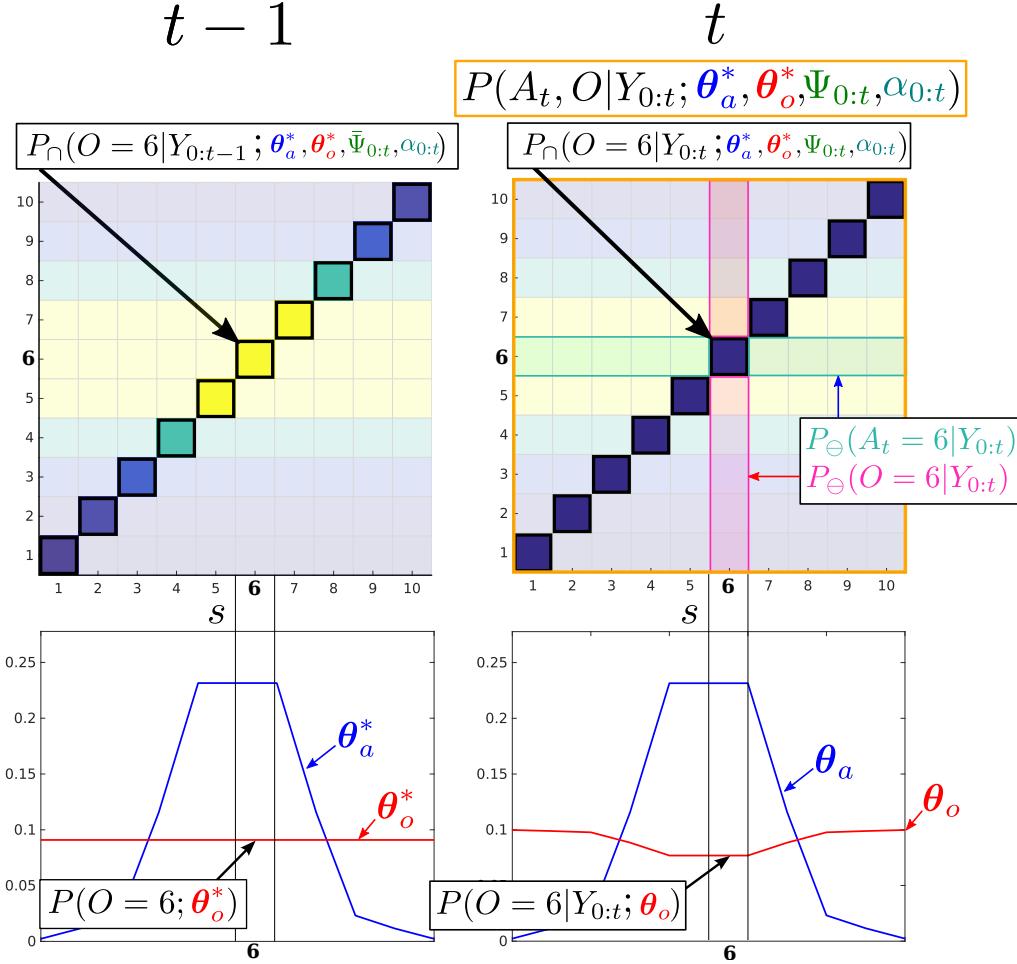


Figure 7. Filtered Marginals. Illustration of the agent and object marginal update, Equation 21. The joint distribution parameters which are independent $A \ominus O$ are pale and the dependent areas $A \cap O$, where $P(Y_t < 1 | A_t, O)$, are bright. MLMF only evaluates the joint distribution in dependent states. For each state s of the marginals $1, \dots, 10$ the difference of the marginals inside the dependent area, before and after the measurement likelihood is applied, is evaluated and removed from the marginals $P(A_t | Y_{0:t-1}, u_{1:t}; \theta_a)$, $P(O | Y_{0:t-1}; \theta_o)$ leading to $\bar{P}(A_t | Y_{0:t}; \theta_a)$, $\bar{P}(O | Y_{0:t}; \theta_o)$ (we did not show $u_{1:t}$ for ease of notation).

functions	parameters	description
$\bar{P}(A_t Y_{0:t}, u_{1:t})$: θ_a	filtered marginals
$\bar{P}(O Y_{0:t})$: θ_o	
$P(A_t u_{1:t})$: θ_a^*	joint marginals
$P(O)$: θ_o^*	
$P(Y_{0:t} u_{1:t})$: $\alpha_{0:t} \in \mathbb{R}$	evidence
$P(Y_{0:t} A_t, O, u_{1:t})$: $\Psi_{0:t} = \{(Y_i, l_i)\}_{i=0:t}$	likelihood history

Table 1. MLMF functions with associated parameters. The marginal parameters are the discretisation of the state space $\theta \in \mathbb{R}^N$, $\theta^{(s)}$ correspond to the probability being in state s .

4 SPACE & TIME COMPLEXITY

342 For discussion purposes we consider the case of three beliefs, namely that of the agent and two other
 343 objects $O^{(1)}$ and $O^{(2)}$, we subsequently generalise. As stated previously M stands for the number of filtered

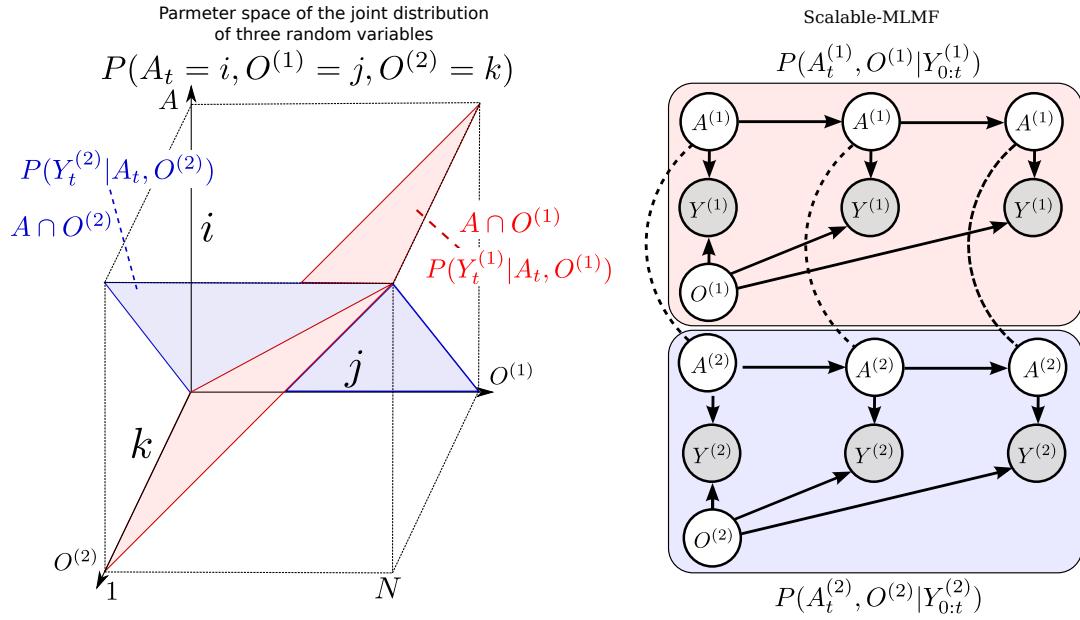


Figure 8. *Left:* Joint distribution $P(A, O^{(1)}, O^{(2)})$ of the agent and two objects ($Y_{0:t}$ and $u_{1:t}$ omitted). Each likelihood function, $P(Y|A, O^{(1)})$, $P(Y|A, O^{(2)})$ corresponds to a hyperplane in the joint distribution. The state space is discretised to N bins giving a potential total of N^3 parameters for the joint distribution (Histogram case). *Right:* Scalable-MLMF Each agent-object joint distribution pair is modelled independently. For clarity we have left out the action random variable u which is linked to every agent node. Two joint distributions $P(A^{(1)}, O^{(1)} | Y_{0:t}^{(1)})$ and $P(A^{(2)}, O^{(2)} | Y_{0:t}^{(2)})$ parametrise the graphical model. The dashed undirected lines represent a wanted dependency, if present $O^{(1)}$ and $O^{(2)}$ are to be dependent through A . In the standard setting there will be no exchange of information between the individual joint distributions. However we demonstrate later on how we perform a one time transfer of information when one of the objects is sensed.

344 random variables including the agent and N is the number of discrete states in the world. In the following
 345 section, we compare the space and time complexity of MLMF-SLAM with Histogram-SLAM.

346 4.1 Space complexity

347 Figure 8 *Left* illustrates the volume occupied by the joint distribution for a space with N states. Histogram-
 348 SLAM would require N^3 parameters for the joint distribution $P(A, O^{(1)}, O^{(2)})$ and N^M parameters for M
 349 random variables.

350 For MLMF-SLAM, each random variable requires two sets of parameters, θ and θ^* (see Table 1). Given
 351 M random variables, the initial number of parameters is $M(2N)$. At every time step the likelihood memory
 352 function increments by one measurement and offset, $(Y_t, l = 0)$ (Algorithm 1). Given a state space of size
 353 N , there can be no more than N different measurement functions (one for each state). In the worst case
 354 scenario the number of memory likelihood function parameters $\Psi_{0:t}$, Equation 15, will be N . The total
 355 number of parameters is $M(2N) + N$ which gives a final worst case space complexity linear in the number
 356 of random variables, $\mathcal{O}(NM)$.

357 4.2 Time complexity

358 For Histogram-SLAM, the computational cost is equivalent to that of the space complexity, $\mathcal{O}(N^M)$,
 359 since every state in the joint distribution has to be summed to obtain all the marginals.

360 For MLMF-SLAM, every state in the joint distribution's state space which has been changed by the
 361 likelihood function has to be summed, see Figure 7 on page 16. As a result the computational complexity
 362 is directly related to the number of dependent states $|A \cap O|$. In Figure 7, this corresponds to states where
 363 $i = j$ and there are N out of a total N^2 states for that joint distribution. Figure 8 *Left* illustrates a joint
 364 distribution with N^3 states. The dependent states $|A \cap O^{(1)} \cap O^{(2)}|$ are those which are within the blue and
 365 red planes (where the likelihood evaluates to zero) and comprise N^2 states each, giving a total of $2N^2 - N$
 366 dependent states (negative is to remove the states we count twice at the intersection of the blue and red
 367 plane).

368 The likelihood term $P(Y_t|A_t, O^{(1)})$ evaluates states to zero which satisfy $i = j \forall k$, as the
 369 measurement of object $O^{(1)}$ is independent of object $O^{(2)}$. With 3 objects, the joint distribution would
 370 be $P(A_t = i, O^{(1)} = j, O^{(2)} = k, O^{(l)} = l)$ then the likelihood $P(Y_t|A_t, O^{(1)})$ evaluated to zero for $i = j$
 371 $\forall k, l$ which would mean N^3 dependent states. In general, for M random variables the computational cost is
 372 $(M-1)N^{M-1}$ which gives $\mathcal{O}(N^{M-1})$ as opposed to the Histogram-SLAM's $\mathcal{O}(N^M)$. The computation
 373 complexity in this setup is still exponential but to the order $M-1$ as opposed to M which nevertheless
 374 quickly limits the scalability as more objects are added.

375 Computing the value of a dependent state (i, j, k) in the joint distribution required evaluating Equation
 376 12 which contains a product of N likelihood functions, in the worst case scenario. However the likelihood
 377 functions are not overlapping and binary. As a result the complete product does not have to be evaluated
 378 since only one likelihood function will effect the state (i, j, k) . Thus evaluating Equation 12 yields a cost
 379 of $\mathcal{O}(1)$ and **not** $\mathcal{O}(N)$.

5 SCALABLE EXTENSION TO MULTIPLE OBJECTS

380 To make the MLMF filter scalable we introduce an **independence assumption** between the objects and
 381 model the joint distribution (Equation 23) as a product of agent-object joint distributions:

$$P(A_t, O^{(1)}, \dots, O^{(M-1)} | Y_{0:t}, u_{1:t}) = \prod_{i=1}^{M-1} P(A_t^{(i)}, O^{(i)} | Y_{0:t}^{(i)}, u_{1:t}) \quad (23)$$

382 The measurement variable Y_t , is the vector of all agent-object measurements, $Y_t = [Y_t^{(1)}, \dots, Y_t^{(M-1)}]^T$.
 383 Each agent-object joint distribution has its own parametrisation of the agent's marginal, $A_t^{(1)}, \dots, A_t^{(M-1)}$
 384 which combine to give the overall marginal of the agent A_t . The computation of each object marginal
 385 $P(O^{(i)} | Y_{0:t}^{(i)})$ is independent of the other objects. This is evident from the marginalisation see Equation
 386 24-25.

$$P(O^{(i)} | Y_{0:t}^{(i)}, u_{1:t}) = \sum_{A_t^{(i)}} P(A_t^{(i)}, O^{(i)} | Y_{0:t}^{(i)}, u_{1:t}) \quad (24)$$

$$P(A_t | Y_{0:t}, u_{1:t}, u_{1:t}) = \prod_{i=1}^{M-1} P(A_t^{(i)} | Y_{0:t}^{(i)}, u_{1:t}) \quad (25)$$

	space	time
Histogram	$\mathcal{O}(N^M)$	$\mathcal{O}(N^M)$
MLMF	$\mathcal{O}(MN)$	$\mathcal{O}(N^{(M-1)})$
scalable-MLMF	$\mathcal{O}(MN)$	$\mathcal{O}(MN)$

Table 2. Time and space complexity summary For both MLMF and scalable-MLMF the worst case scenario is reported for the space complexity.

387 The independence assumption will create an unwanted effect with respect to agent's marginal
 388 $P(A_t|Y_{0:t}, u_{1:t})$. At initialisation the agent marginals should be equal, $P(A_0|Y_0) = P(A_0^{(i)}|Y_0^{(i)}) \forall i$,
 389 however this is not the case because of Equation 25. To overcome this we define the final marginal,
 390 $P(A_t|Y_{0:t}, u_{1:t})$, of the agent as being the average of all the individual pairs $P(A_t^{(i)}|Y_{0:t}^{(i)}, u_{1:t})$.

$$P(A_t|Y_{0:t}, u_{1:t}) := \frac{1}{M-1} \sum_{i=1}^{M-1} P(A_t^{(i)}|Y_{0:t}^{(i)}, u_{1:t}) \quad (26)$$

391 Figure 8 (*Right*), depicts the graphical model of the scalable formulation. As each joint distribution pair
 392 has its own parametrisation of the agent's marginal and these do not subsequently get updated by one
 393 another, the information gained by one joint distribution pair is **not transferred**. A solution is to transfer
 394 information between the marginals $A^{(i)}$ at specific intervals namely when one of the objects is sensed by
 395 the agent.

396 The exchange of information of one joint distribution to another is achieved through the agent's marginals
 397 $A^{(i)}$ according to Algorithm 3. The measurement update is the same as previously described in Algorithm
 398 2 in the case of no positive measurements of the objects. If the agent senses an object, all of the agent
 399 marginals of the remaining joint distributions are set to the marginal of the joint distribution pair belonging
 400 to the positive measurement $Y_t^{(i)}$.

401 Figure 9, depicts the process of information exchange between object $O^{(1)}$ and $O^{(2)}$ in the event
 402 that the agent senses $O^{(2)}$. Prior to the positive detection, both marginals $P(A_t^{(1)}|Y_{0:t-1}^{(1)}, u_{1:t})$ and
 403 $P(A_t^{(2)}|Y_{0:t-1}^{(2)}, u_{1:t})$ occupy the same region and are identical. When the agent senses $O^{(2)}$ the line defined
 404 by the measurement likelihood function $P(Y_t^{(2)}|A_t^{(2)}, O^{(2)})$ becomes a hard constraint implying that both
 405 the agent and $O^{(2)}$ have to satisfy this constraint. Figure 10 shows marginals at initialisation, prior contact
 406 between the agent and object and the after the measurement (post contact) has been integrated into the
 407 marginals (resulting from the joint distributions in Figure 9).

408 The result of introducing a dependency between the objects through the agent's marginals in the event of
 409 a sensing and treating them independently gives the same solution as the histogram filter in this particular
 410 case. However as each individual marginal $A_t^{(i)}$ diverges from the other marginals, the filtered solution
 411 will diverge from the histogram's solution. We assume however that the objects are weakly dependent
 412 and sharing information during positive sensing events is sufficient. In section 6.2 we will evaluate the
 413 independence assumption with respect to the histogram filter.

414 Table 2 summarises the time and space complexity for the three filters.

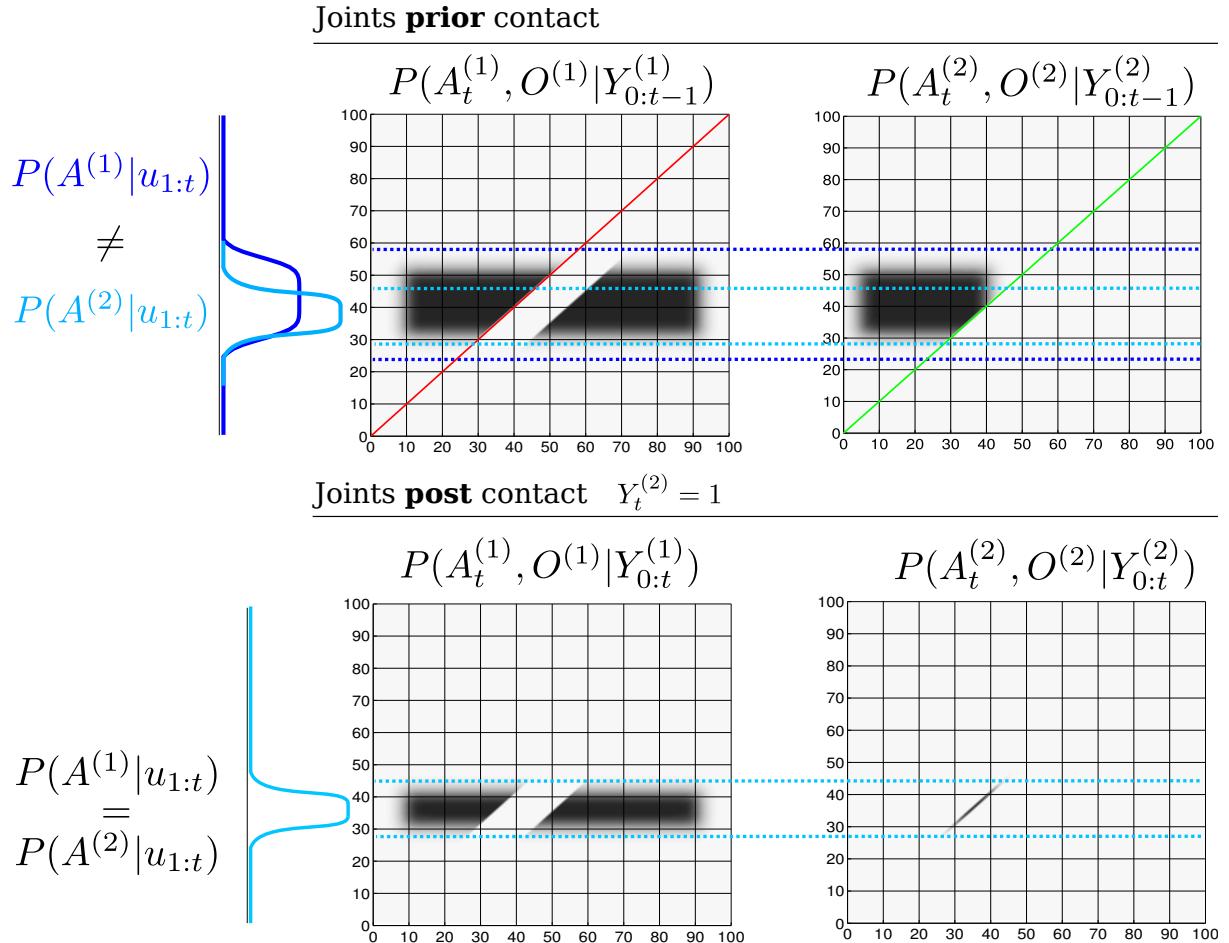


Figure 9. Transfer of information (joint distributions) *Top:* Joint distributions of $P(A_t^{(1)}, O^{(1)}|Y^{(1)})$ and $P(A_t^{(2)}, O^{(2)}|Y^{(2)})$ prior sensing, $Y_t^{(2)} = 1$, see Figure 10 (*Top right*) for the corresponding marginals. The red and green lines across the joint distributions correspond to the region in which the likelihood functions $P(Y_t^{(1)}|A_t^{(1)}, O^{(1)})$ and $P(Y_t^{(2)}|A_t^{(2)}, O^{(2)})$ will change the joint distributions. The dotted blue lines are to ease the comparison of the joint distributions prior and post sensing. *Bottom right:* After the agent has sensed $O^{(2)}$, all the probability mass which was not overlapping the green line becomes an infeasible solution to the agent and object locations. At this point the marginals $P(A_t^{(1)}|u_{1:t}) \neq P(A_t^{(2)}|u_{1:t})$ are no longer equal (see the blue marginals *Top*). *Bottom left:* The constraint imposed by the likelihood function of the second object (green line) is transferred to the joint distribution of the first object according to Algorithm 3. This results in a change in the joint distribution $P(A_t^{(1)}, O^{(1)}|Y^{(1)})$, which satisfies the constraints imposed by the agent's marginal from the joint distribution $P(A_t^{(2)}, O^{(2)}|Y^{(2)})$.

6 EVALUATION

415 We conduct three different types of evaluation to quantify the scalability and correctness of the scalable-
 416 MLMF filter. The first experiment tests the scalability of our filter in terms of processing time taken per
 417 motion-measurement update cycle. The second experiment evaluates the independence assumption made
 418 in the scalable-MLMF filter between the objects. The third and final experiment determines the effect of
 419 the memory size on a search policy to locate all the objects in the *Table* world.

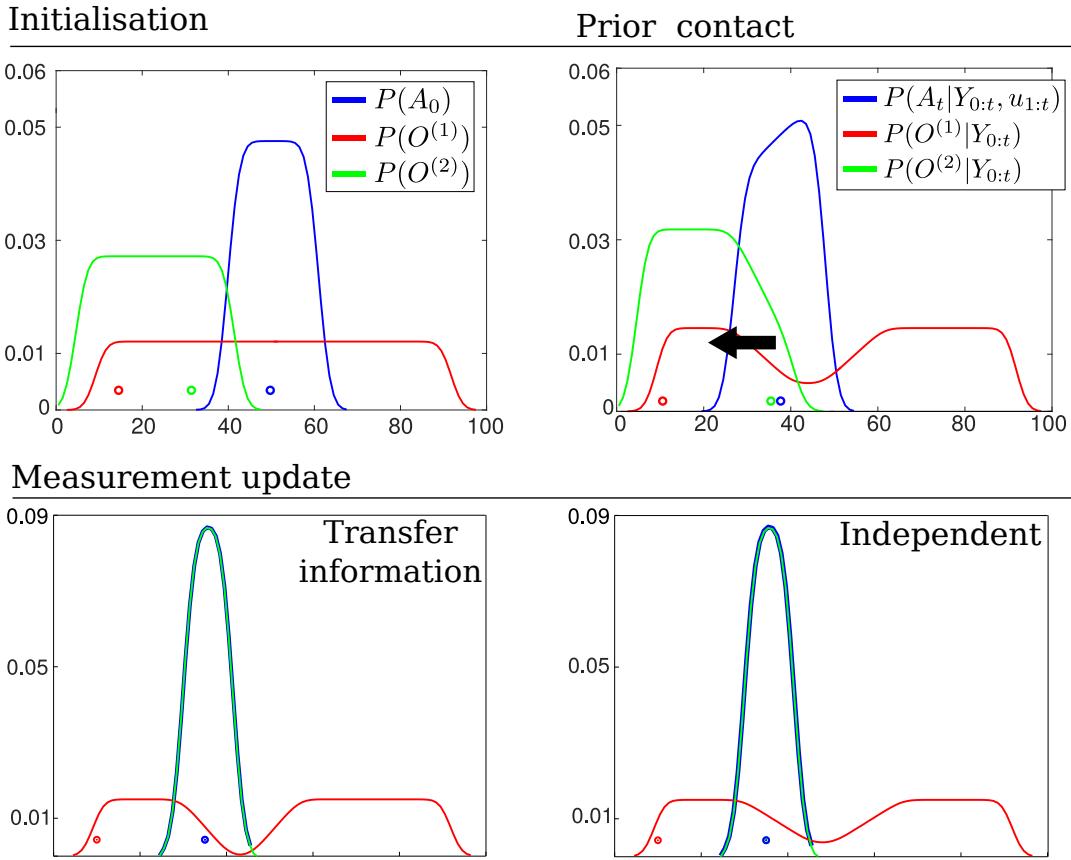


Figure 10. Transfer of information (marginals) *Top left:* Initial beliefs of the agent and object’s location. The agent moves to the left until it senses object $O^{(2)}$. *Top right:* Marginals prior the agent entering in contact with the green object, see Figure 9 (Top) for an illustrate of the joint distributions. *Bottom left:* resulting marginals after setting the agent marginals of each joint distribution equal $A_t^{(1)} = A_t^{(2)}$ according to Algorithm 3. The object marginal $P(O^{(2)}|Y_{0:t})$ is recomputed. *Bottom Right:* resulting marginals in which the objects have no influence on one another. Note that a transfer of information has caused a change in the marginal $O^{(1)}$.

420 6.1 Evaluation of time complexity

421 We measured the time taken by the motion-measurement update loop, as a function of the number
 422 beliefs and number of states per belief. We started with a 100 states per belief and gradually increase it
 423 to 10’000’000 over 50 steps. Each of the 50 steps treated 2 to 25 objects. Figure 11 *left* illustrates the
 424 computational cost as a function of number of states and objects. For each state-object pair 100 motion-
 425 measurement updates were performed. Most of the trials returned time updates below 1 Hz. Figure 11 *right*
 426 shows the computational cost as a function of the number of states plotted for 6 different filter runs with a
 427 different number of objects. As the number of states increases exponentially so does the computational
 428 cost. Note the cost increases at the same rate as the number of states meaning that the computational
 429 complexity is linear with respect to the number of states. This result is in agreement with the asymptotic
 430 time complexity.

431 6.2 Evaluation of the independence assumption

432 In section 5 we made the assumption (for scalability reasons) that the objects’ beliefs are independent
 433 of one another. This assumption is validated by comparing the MLMF filter on three random variables,

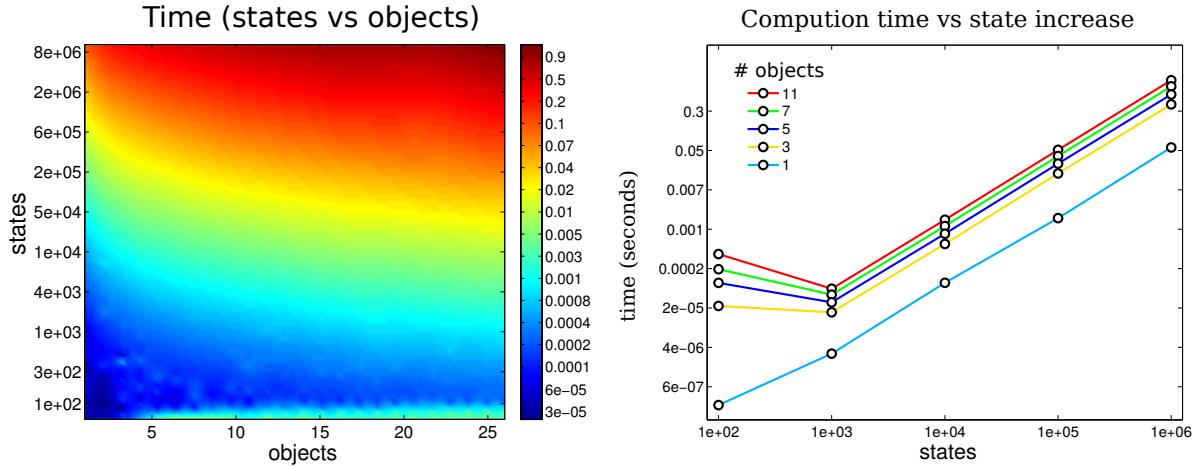


Figure 11. Time complexity: *left*: mean time taken for a loop update (motion and measurement) as a function of the number of states in a marginal and the number of objects present. *right*: time taken for a loop update with respect to the number of states in the marginal. The colour coded lines are associated with the number of objects present. The computational cost is plotted on a log scale. As the number of states increases exponentially the computational cost matches it.

434 an agent and two objects, with the ground truth which we obtain from the standard histogram filter. For
 435 each of the three beliefs (the agent and two objects), 100 different marginals were generated and the true
 436 locations (actual position of the agent and objects) were sampled. Figure 12 *Top-left* illustrates one instance
 437 of the initialisation of the agent and object marginals with their associated sampled true position. The agent
 438 carries out a sweep of the state space for each of the marginals and the policy is saved and run with the
 439 scalable-MLMF filter. In the first experiment we assumed that the objects are completely independent and
 440 that there was no transfer of information between the pair-wise joint distributions. In the second and third
 441 experiments there is an exchange of information as described in Algorithm 3. Here we compare the effect
 442 of using the product of the agent's marginals, Equation 25, with the average of the marginals, Equation 26.
 443 We expect the average of the the agent's marginal to yield a result closer to the ground truth as the marginal
 444 of the agent $P(A_t|Y_{0:t}, u_{1:t})$ at initialisation is the same as the ground truth (the Histogram-SLAM's). As
 445 for the marginal of the objects $P(O^{(i)}|Y_{0:t})$ we expect the difference between them to be independent of
 446 whether the product or average of the agent's marginal is used. This results from Algorithm 3. When an
 447 object i is sensed all the corresponding agent marginals $P(A^{(j)}|u_{1:t})$ are set equal to $P(A^{(i)}|u_{1:t})$ and not
 448 to $P(A_t|Y_{0:t}, u_{1:t})$. This is a design decision of our information transfer heuristic. There are many other
 449 possibilities but this is one of the simplest. For each of the 100 sweeps the ground truth is compared with
 450 the scalabe-MLMF using the Hellinger distance (Equation 27)

$$H(P, Q) = \frac{1}{\sqrt{2}} \|\sqrt{P} - \sqrt{Q}\|_2 \quad (27)$$

451 which is a metric which measures the distance between two probability distributions. Its value lies strictly
 452 between 0 (the two distributions are identical) and 1 (no overlap between them). Figure 12 shows the
 453 kernel density distribution of the Hellinger distances taken at each time step for all 100 sweeps. In the
 454 *Top-left* of the figure, for the case when no transfer of information is applied, all the marginals are far
 455 from the ground truth. This results from the introduction of the independence assumption, necessary to
 456 scale the MLMF. Figure 12 *Bottom* shows the results for difference between the product and average of

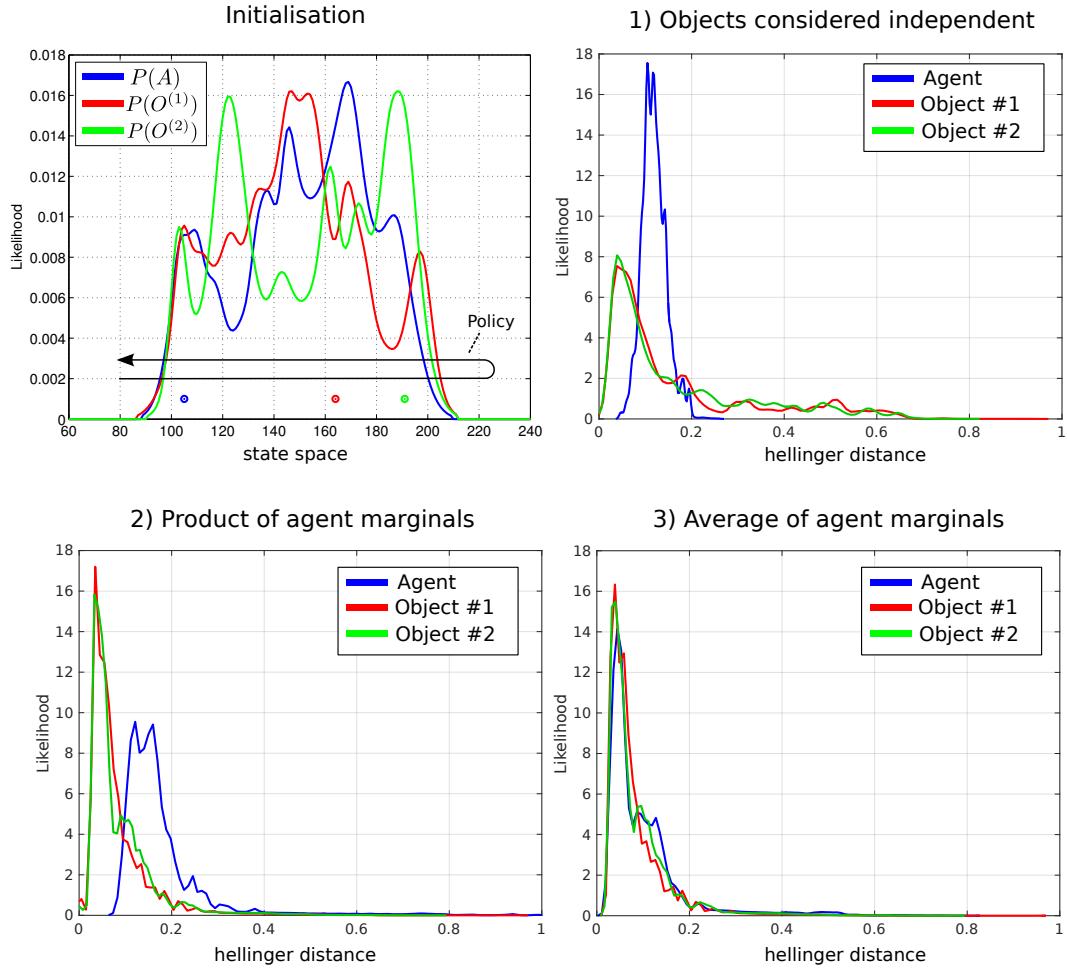


Figure 12. Comparison of scalable-MLMF and the histogram filter A deterministic sweep policy was carried out for 100 different initialisations of the agent and object beliefs. **Top left:** One particular Initialisation of the agent and object random variables. The true position of the agent and objects were sampled at random. The black arrow indicates the general policy which was followed for each of the 100 sweeps. These were performed for **1)** scalable-MLMF with objects considered to be independent at all times (no Algorithm 3). **2)** Agent marginal $P(A_t|Y_{0:t}, u_{1:t})$ is the product of marginals $P(A_t^{(i)}|Y_{0:t}^{(i)}, u_{1:t})$, Equation 25. **3)** marginal $P(A_t|Y_{0:t}, u_{1:t})$ is taken to be the average of all marginals $P(A_t^{(i)}|Y_{0:t}^{(i)}, u_{1:t})$, Equation 26. For each of these three experiment we report the kernel density estimation over the Hellinger distances taken at every time step between ground truth (from histogram filter) and scalable-MLMF.

457 the agents marginals. As expected there is no difference between the objects' marginals when considering
 458 both methods (product and average) with respect to the ground truth. The predominant difference occurs in
 459 the agent's marginal $P(A_t|Y_{0:t}, u_{1:t})$. This is also expected and prompted the introduction of the average
 460 method instead of the product.

461 The scalable-MLMF information exchange heuristic will not lead to any of the objects marginals
 462 probability mass being falsely removed during the information transfer, which is close to a winner-take-all
 463 approach in terms of beliefs. When object i is sensed its associated agent marginal is set to all other
 464 agent-object joint pairs, which results in the information accumulated in the j th agent marginals being
 465 replaced by the i th.

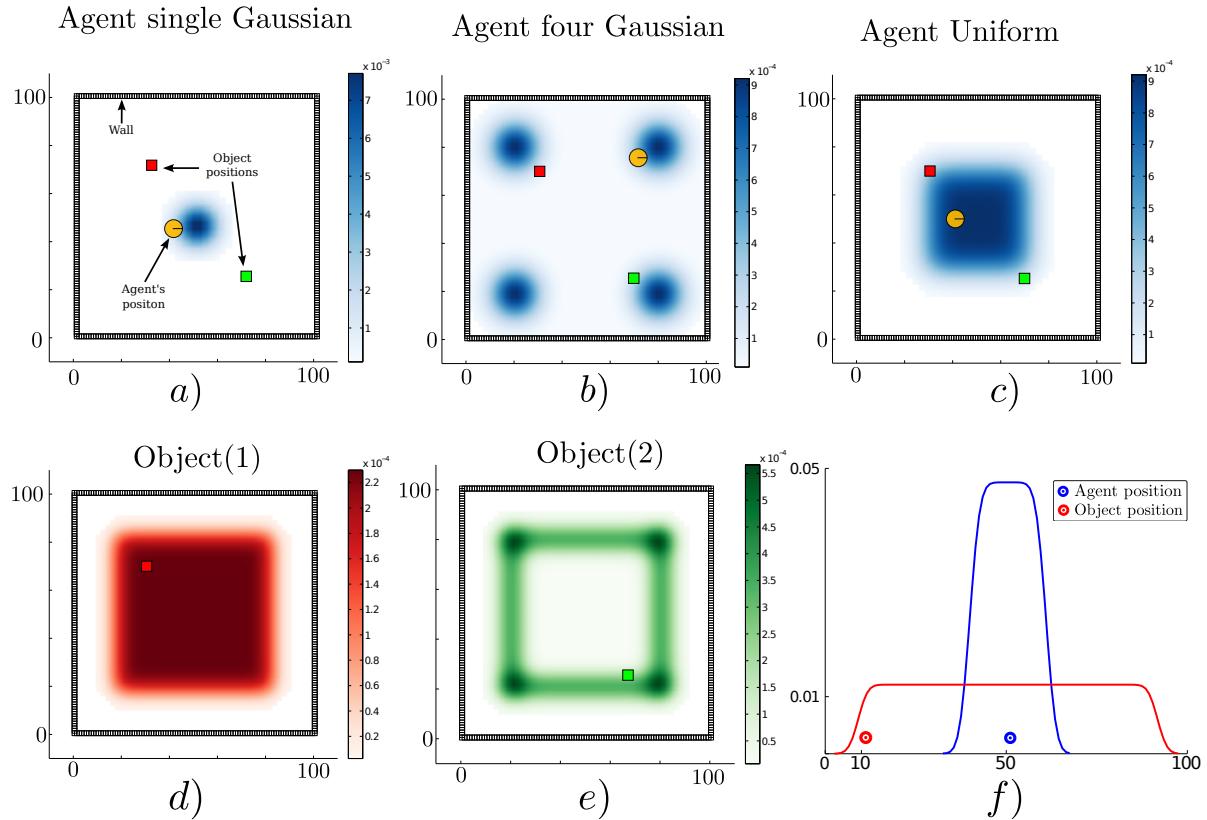


Figure 13. Agent's prior beliefs. Two types of environment, the first is a 2D world where the agent lives in a square surrounded by a wall whilst the second is a 1D world. In the 2D figures the agent is illustrated by a circle with a bar to indicate its heading. The true location of the objects are represented by colour coded squares. *Top row* three different initialisations of the agent's location. *Bottom row* d) the agent's prior beliefs with respect to the location of the first object and e) belief of the second object's location. *bottom row* f) 1D world with one object.

466 6.3 Evaluation of memory

467 The memory measurement likelihood function $P(Y_{0:t}|A_t, O, u_{1:t}; \Psi_{0:t})$ is parameterised by the history
 468 of all the measurement likelihood functions which have been applied on the joint distribution since
 469 initialisation. As detailed previously there can be no more than $|\Psi_{0:t}| \leq N$ different measurement
 470 likelihood functions added to memory. In the case of a very large state space this might be cumbersome.
 471 We investigate how restricting the memory size, the number of parameters $|\Psi_{0:t}|$, can impact on the
 472 decision process in an Active-SLAM setting. Given our set up a breadth-first search in the action space is
 473 chosen with a one time step horizon, making it a greedy algorithm. The objective function utilised is the
 474 information gain of the beliefs after applying an action, Equation 28.

$$u_t = \arg \max_{u_t} H\{P(A_{t-1}, O|Y_{0:t-1}, u_{1:t-1})\} - \mathbb{E}_{Y_t} [H\{P(A_t, O|Y_{0:t}, u_{1:t})\}] \quad (28)$$

475 For each action the filter is run forward in time and all future measurements since we cannot know ahead
 476 of time the actual measurement. The information gain is the difference between the current entropy (defined
 477 by $H\{\cdot\}$) and the future entropy after the simulated motion and measurement update. The action with the
 478 highest information gain is subsequently selected. This is repeated at each time step. Figure 13 illustrates
 479 the environment setup for a 1D and 2D case. The agent's task is to find the objects in the environment.

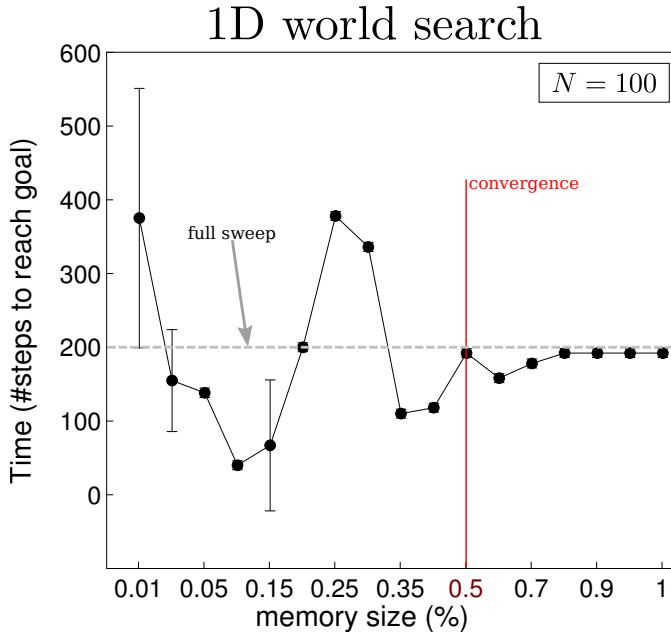


Figure 14. Memory size vs time to find object in 1D Results of the effect of the memory size on the decision process for the 1D search illustrated in Figure 13 f). The memory size is reported as the percentage of total number of states present in the marginal space. At 100% the size of the memory is equal to that of the state space, $N = 100$ in this case. A total sweep of the entire state space would result in a total of 200 steps, the dotted grey line in the above figure. When no restrictions are placed on the memory size the policy following the greedy approach takes around 180 steps. This result converges when the number of parameters $|\Psi_{0:t}|$ of the memory likelihood function is greater than 50% of the original state space.

480 For the 2D search we consider three different initialisations (single-Gaussian, four-Gaussian, Uniform)
 481 for the agent's belief where there are two objects to be found. Ten searches are carried out for each of the
 482 three initialisations of the agent's beliefs. The agent's true location, for each search, is sampled from its
 483 initial belief, and the objects' locations (red and green squares in Figure 13) are kept fixed throughout all
 484 searches. Each search is repeated for 18 different memory sizes ranging from 1 to N (the number of states).
 485 For the 1D search case one object is considered since adding more objects makes the search easier and the
 486 interest lies in the memory effects of the search and not the search itself. In Figures 14-15 we report on the
 487 time taken to find all objects with respect to a given memory size which is shown as the percentage of the
 488 total number of states. In the 1D search case the time variability taken to find the object converges when
 489 the memory size is at 60% of the original state space. As for the 2D search with 2 beliefs (agent & 1 object)
 490 the convergence depends on the agent's initial belief. For the 1-Gaussian (green line) all searches take
 491 approximately the same amount of time after a memory size of 9%. As for the remaining two initialisations
 492 convergence is achieved at 48%. The same holds true for the case of 3 beliefs (agent & 2 objects).

493 In the 2D searches, the memory size has a less impact on the time taken to find the objects than in the 1D
 494 (which is a special search case). Only when the memory size is less than 6% is there a significant change.
 495 We conclude that at least in the case of the greedy one step-look ahead planner which is frequently used
 496 in the literature, the size of the memory seems not to be a limiting factor in terms of the time taken to
 497 accomplish the search.

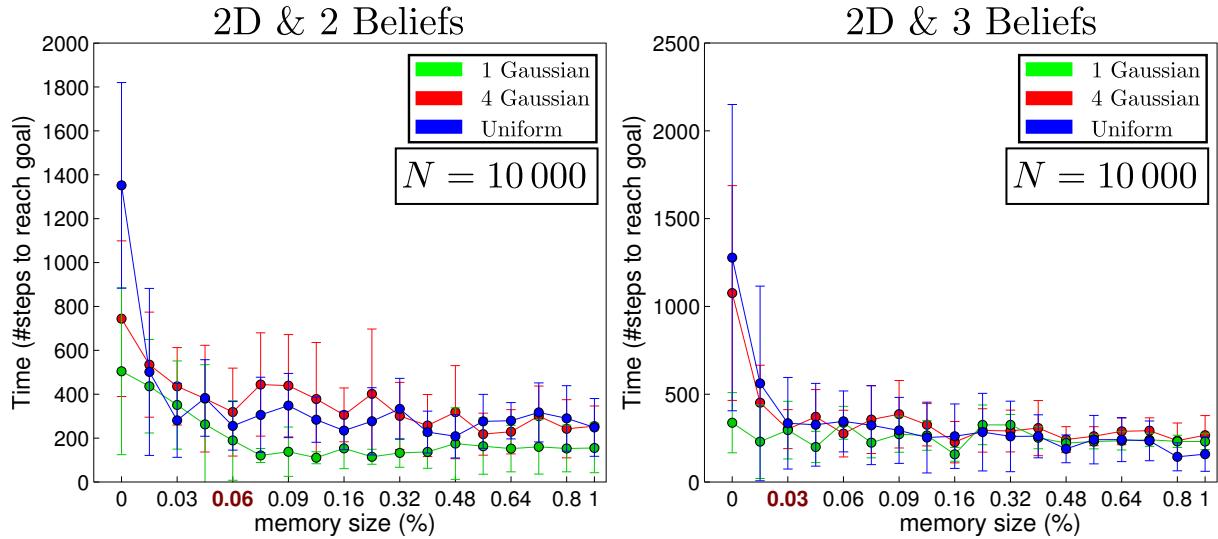


Figure 15. Memory size vs time to find objects in 2D. The initial beliefs correspond to those of Figure 13, a) for Gaussian (green line), b) 4 Gaussians (red line) and c) Uniform (blue line), both objects are initialised according to d) and e).

7 CONCLUSION

498 This work addresses the Active-SLAM filtering problem for scenarios in which sensory information
 499 relating to the map is very limited. Current SLAM algorithms filter the errors originating from sensory
 500 measurements and not prior uncertainty. By making the assumption that the joint distribution of all the
 501 random variables is a multivariate Gaussian, inference is tractable. Since the origin of the uncertainty does
 502 not originate from the measurement noise, no assumption can be made about the structure of the joint
 503 distribution. In this case a suitable filter would be the histogram which makes no assumption about the
 504 shape or form taken by the joint distribution. However, the space and time complexity are exponential with
 505 respect to the number random variables and this is a major limiting factor for scalability.

506 The main contribution of this work is a formulation of a histogram Bayesian state space estimator in
 507 which the computational complexity is both linear in time and space. A different approach to other SLAM
 508 formulations has been taken in the sense that the joint distribution is not explicitly parameterised avoiding the
 509 exponential increase in parameter space which would otherwise have been the case. The MLMF parameters
 510 consist of the marginals and the history of measurement functions which have been applied. By solely
 511 evaluating the joint distribution at the states which are affected by the current measurement function whilst
 512 taking into account the memory, the MLMF filter obtains the same filtered marginals as the histogram
 513 filter. Further, the worst case space complexity is linear rather than exponential and the time complexity
 514 remains exponential but increases at lower rate than in the histogram filter. In striving to make the filter
 515 scalable we make the assumption that the objects are independent. An individual MLMF is used for each
 516 agent-object pair. We evaluate the difference between the scalable-MLMF with a ground truth provided by
 517 the histogram filter for 100 different searches with respect to the Hellinger distance. We conclude that the
 518 divergence is relatively small and thus the scalable-MLMF filter provides a good approximation to the true
 519 filtered marginals. We evaluate the time taken to perform a motion-update loop for different discretisations
 520 of the state space (100 to 10'000'000 states) and number of objects (2 to 25). In most of the cases we
 521 achieve an update cycle rate below 1Hz. We evaluate how the increase of the number of states effects the
 522 computational cost and find the relationship to be linear and thus in agreement with our analysis of the

523 asymptotic growth rate. We analyse the effect of the memory size (the remembered number of measurement
524 likelihood functions) on the decision theoretic process of reducing the uncertainty of the map and agent
525 during a search task. We conclude that in the 2D case the memory size has much less effect than in the 1D
526 case and that it is unnecessary to remember every single measurement function.

527 This implies that the MLMF and scalable-MLMF that we have are a computationally tractable means
528 of performing SLAM in a case scenario in which mostly negative information is present and the joint
529 distribution cannot be assumed to have any specific structure. Furthermore, the filter can be used at a higher
530 cognitive level than the processing of raw sensory information as is often the case in Active-SLAM. MLMF
531 would be well suited for reasoning tasks where the robot's field of view is limited.

532 An interesting future extension could be to make the original MLMF filter scalable without introducing
533 assumptions. One possibility could be to consider Monte Carlo integration methods for inference. These
534 can scale well to high dimensional spaces whilst still providing reliable estimates. A second possibility
535 could be to investigate the use of Gaussian Mixtures as a form of parameterisation of the marginals to blend
536 our filter with EKF-SLAM. This would allow the parameters to grow quadratically with respect to the
537 dimension of the marginal space as opposed to exponentially as is the case with the histogram and MLMF
538 filters.

8 APPENDIX

539 8.1 MLMF Algorithm

Algorithm 2: MLMF-SLAM

input :

measurements

$\mathbf{Y}_t, \mathbf{u}_t$

joint distribution parameters:

$P(A_{t-1}|u_{1:t-1}) P(O), \Psi_{0:t-1}, \alpha_{0:t-1}$

filtered marginals:

$P(A_{t-1}|Y_{0:t-1}, u_{1:t-1}), P(O|Y_{0:t-1})$

output:

joint parameters:

$P(A_t|u_{1:t}), \Psi_{0:t}, \alpha_{0:t}$

filtered marginals:

$P(A_t|Y_{0:t}, u_{1:t}), P(O|Y_{0:t})$

initialisation

$$P(A_0; \theta_a) := P(A_0; \theta_a^*)$$

$$P(O; \theta_o) := P(O; \theta_o^*)$$

$$\Psi_0 := \{\}$$

$$\alpha_0 := 0$$

540

motion update

$$P(A_t|u_{1:t}) = \sum_{A_{t-1}} P(A_t|A_{t-1}, \mathbf{u}_t) P(A_{t-1}|u_{1:t-1})$$

$$P(A_t|Y_{0:t-1}, u_{1:t}) = \sum_{A_{t-1}} P(A_t|A_{t-1}, \mathbf{u}_t) P(A_{t-1}|Y_{0:t-1}, u_{1:t-1})$$

$$\bar{\Psi}_{0:t} \leftarrow \Psi_{0:t-1} : \text{Algorithm 1 (motion update)}$$

measurement update

$$\alpha_{0:t} = \alpha_{0:t-1} + \sum_{A_t} \sum_O \left(P(\mathbf{Y}_t|A_t, O) - 1 \right) P_{\cap}(A_t, O, Y_{0:t-1}|u_{1:t})$$

$$P(Y_{0:t}|u_{1:t}) = 1 + \alpha_{0:t}$$

$$P(A_t|Y_{0:t}) = P(A_t|Y_{0:t-1}) - \left(P_{\cap}(A_t|Y_{0:t-1}) - P_{\cap}(A_t|Y_{0:t}) \right)$$

$$P(O|Y_{0:t}) = P(O|Y_{0:t-1}) - \left(P_{\cap}(O|Y_{0:t-1}) - P_{\cap}(O|Y_{0:t}) \right)$$

$$\Psi_{0:t} \leftarrow \bar{\Psi}_{0:t} : \text{Algorithm 1 (measurement update)}$$

541 8.2 Scalabe-MLMF Algorithm

Algorithm 3: Scalable-MLMF: Measurement Update

input : $P(A_t^{(i)}|u_{1:t}), P(A_t^{(i)}|Y_{0:t-1}^{(i)}, u_{1:t})$

$$P(O^{(i)}), P(O^{(i)}|Y_{0:t-1}^{(i)}, u_{1:t})$$

$$Y_t^{(i)}$$

$i =$

$v = 1, \dots, IV$

▷ If object i has been sensed by the agent
 (i)

1 if $Y_t^{(i)} == 1$ **then**

$$P(O^{(i)}|Y_{0:t}^{(i)}) \leftarrow P(O^{(i)}|Y_{0:t-1}^{(i)}) ; \quad \triangleright \text{ measurement update Algo. 2}$$

542 **3** $P(A_t^{(i)} | Y_{0:t}^{(i)}, u_{1:t}) \leftarrow P(A_t^{(i)} | Y_{0:t-1}^{(i)}, u_{1:t})$
4 **forall the $i \in \{1, \dots, M-1\}\}$ do**

4 **forall** the $j \in (1, \dots M - 1) \setminus i$ **do**

$$P(A_t^{(j)}|Y_{0:t}, u_{1:t}) = P(A_t^{(i)}|Y_{0:t}, u_{1:t})$$

$$P(A_t^{(j)}|u_{1:t}) = P(A_t^{(i)}|u_{1:t})$$

$$P(O^{(j)}|Y_{0:t}^{(i)}) \leftarrow \sum P(A_t^{(j)}, O^{(j)}|Y_{0:t}^{(i)})$$

$$I(\mathcal{O}^+ | \mathcal{I}_{0:t}) \leq \sum_{A^{(j)}} I(A_t^+, \mathcal{O}^+ | \mathcal{I}_{0:t})$$

8 else

forall the $i \in (1, \dots M)$ **do**

measurement update Algo. 2

543 8.3 Recursion example

544 Derivation of the filtered joint distribution, $P(A_t, O, Y_t | Y_{0:t}, u_{1:t})$, for two updates. At initialisation when
 545 no action has yet been taken the filtered joint distribution is the product of the initial marginals and first
 546 likelihood function:

$$P(A_0, O, Y_0) = P(O)P(A_0)P(Y_0|A_0, O) \quad (29)$$

The first action, u_1 is applied, which to get the filtered joint distribution is marginalised:

$$P(A_1, O, Y_0 | u_1) = P(O) \sum_{A_0} P(A_1 | A_0, u_1) P(A_0) P(Y_0 | A_0, O) \quad (30)$$

$$= P(O) \sum_{A_0} P(A_1, A_0, Y_0 | u_1, O) \quad (31)$$

$$= P(O)P(A_1, Y_0|u_1, O) \quad (32)$$

$$= P(O)P(Y_0|A_1, O, u_1)P(A_1|u_1, \emptyset) \quad (33)$$

$$= P(O)P(Y_0|A_1, O, u_1)P(A_1|u_1) \quad (34)$$

547 From Equation 32 to 33 we used the Chain rule and the cancellation in Equation 33 arise from the
 548 factorisation of the joint distribution, see Figure 2 on page 6, A 's marginal does not depend on O . After the
 549 application of the first action, the filtered joint has the following form:

$$P(A_1, O, Y_0 | u_1) \equiv P(O)P(A_1|u_1)P(Y_0|A_1, O, u_1) \quad (35)$$

A second measurement Y_1 and action u_2 are integrated into the filtered joint distribution:

$$\begin{aligned} P(A_2, O, Y_{0:1}|u_{1:2}) &= P(O) \sum_{A_1} P(A_2|A_1, u_2) P(A_1|u_1) P(Y_0|A_1, O, u_1) P(Y_1|A_1, O) \\ &= P(O) \sum_{A_1} P(A_2, A_1|u_{1:2}) P(Y_{0:1}|A_1, O, u_1) \\ &= P(O) \sum_{A_1} P(A_2, A_1, Y_{0:1}|O, u_{1:2}) \\ &= P(O) P(A_2, Y_{0:1}|O, u_{1:2}) \end{aligned} \quad (36)$$

$$= P(O) P(Y_{0:1}|A_2, O, u_{1:2}) P(A_2|\cancel{O}, u_{1:2}) \quad (37)$$

550 We expand the function $P(Y_{0:1}|A_2, O, u_{1:2})$ to give a sense of how the likelihood function's positions get
 551 as illustrated in Figure 5 on page 10.

$$P(Y_0, Y_1|A_2, O, u_1, u_2) = P(Y_0|\cancel{Y_1}, A_2, O, u_1, u_2) P(Y_1|A_2, O, \cancel{u_1}, u_2) \quad (38)$$

$$= P(Y_0|A_2, O, u_{1:2}) P(Y_1|A_2, O, u_2) \quad (39)$$

552 The first likelihood of measurement Y_0 is dependent on the last two applied actions whilst the likelihood of
 553 Y_1 is dependent on the last action.

554 Repeating the above for $Y_{2:t}$ and $u_{3:t}$ results in:

$$P(A_t, O, Y_{0:t}|u_{1:t}) = P(O) P(A_t|u_{1:t}) \prod_{i=0}^t P(Y_i|A_t, O, u_{i+1:t}) \quad (40)$$

If $t = 3$, $(Y_{0:3}$ and $u_{1:3})$ according to the above equation we would get:

$$\begin{aligned} P(A_3, O, Y_{0:3}|u_{1:3}) &= P(O) P(A_3|u_{1:3}) P(Y_0|A_3, O, u_{1:3}) \\ &\quad P(Y_1|A_3, O, u_{2:3}) \\ &\quad P(Y_2|A_3, O, u_{3:3}) \\ &\quad P(Y_3|A_3, O, \cancel{u_{4:3}}) \end{aligned} \quad (41)$$

555 We introduce some notation rules, first if $(i + 1) > t$ for $u_{(i+1):t}$ then it cancels out since the current
 556 measurement Y_t cannot depend on a future action $u_{(i+1)}$.

557 8.4 Derivation of the evidence

558 The evidence, also known as the marginal likelihood, is the marginalisation of all non measurement
 559 random variables from the filtered joint distribution $P(A_t, O, Y_{0:t}|u_{1:t})$. We detail below how we compute
 560 the evidence in a recursive manner whilst only considering dependent regions of the joint distribution.

561 We start with the **standard** definition of the evidence:

$$P(Y_{0:t}|u_{1:t}) = \sum_{A_t} \sum_O P(A_t, O, Y_{0:t}|u_{1:t}) \in \mathbb{R} \quad (42)$$

If both A_t and O are random variables defined over a discretised state space of N states, the above double integral will sum a total of N^2 states which is the complete state space of the joint distribution $P(A_t, O, Y_{0:t}|u_{1:t}) \propto P(A_t, O|Y_{0:t}, u_{1:t})$, see Figure 6 on page 13 for an illustrate of such a joint distribution. As we are interested in a recursive computation of the evidence, we consider the gradient:

$$\alpha_t = \nabla_{Y_t} P(Y_{0:t}|u_{1:t}) = P(Y_{0:t}|u_{1:t}) - P(Y_{0:t-1}|u_{1:t}) \quad (43)$$

$$\alpha_t = \sum_{A_t} \sum_O P(A_t, O, Y_{0:t}|u_{1:t}) - P(A_t, O, Y_{0:t-1}|u_{1:t}) \quad (44)$$

$$= \sum_{A_t} \sum_O P(Y_t|A_t, O)P(A_t, O, Y_{0:t-1}|u_{1:t}) - P(A_t, O, Y_{0:t-1}|u_{1:t}) \quad (45)$$

$$= \sum_{A_t} \sum_O (P(Y_t|A_t, O) - 1)P(A_t, O, Y_{0:t-1}|u_{1:t}) \quad (46)$$

562 The gradient α_t is the difference in mass before and after the application the likelihood function,
 563 $P(Y_t|A_t, O)$. The above summation, Equation 46, is over the entire joint distribution state space. We
 564 can take advantage of the fact that the likelihood function is sparse and will only affect a small region of
 565 the joint distribution, which we called the dependent states, \cap . The states which are not affected by the
 566 joint distribution will result in a contribution of zero to Equation 46. We rewrite the gradient update in
 567 terms of only the dependent regions:

$$\alpha_t = \sum_{A_t} \sum_O (P(Y_t|A_t, O) - 1)P_{\cap}(A_t, O, Y_{0:t-1}|u_{1:t}) \quad (47)$$

568 Consider the first update of the evidence at time $t = 0$:

$$\alpha_0 = \sum_{A_t} \sum_O (P(Y_0|A_0, O) - 1)P(A_0, O) \quad (48)$$

569 The one in Equation 49 is the original value of the normalisation denominator before any observation is
 570 made and as the initial joint distribution $P(A_0, O)$ is normalised the value of the denominator is one.

$$P(Y_0) = 1 + \alpha_0 \quad (49)$$

571 For the evidence $P(Y_{0:t}|u_{1:t})$ we consider the summation of all the derivatives α_t from time $t = 0$ until t :

$$P(Y_{0:t}|u_{1:t}) = 1 + \sum_{t=0}^T \alpha_t \quad (50)$$

572 **8.5 Derivation of the marginal**

573 The marginal of a random variable is the marginalisation or integration over all other random variables,
 574 $P(A_t, |Y_{0:t}) = \sum_O P(A_t, O|Y_{0:t})$. Below we give a form of this integration which exploits the independent
 575 regions in the joint distribution.

$$P(A_t, |Y_{0:t}) = \mathbf{P}(\mathbf{A}_t | \mathbf{Y}_{0:t-1}) - \left(\mathbf{P}(\mathbf{A}_t | \mathbf{Y}_{0:t-1}) - P(A_t | Y_{0:t}) \right) \quad (51)$$

576 In Equation 51 we add and subtract $P(A_t | Y_{0:t-1})$ and we further split $P(A_t | Y_{0:t-1})$ into independent
 577 and dependent components:

$$P(A_t, |Y_{0:t}) = P(A_t | Y_{0:t-1}) - \left(\underbrace{P_{\cap}(A_t | Y_{0:t-1}) + P_{\ominus}(A_t | Y_{0:t-1})}_{P(A_t | Y_{0:t-1})} - \underbrace{P_{\cap}(A_t | Y_{0:t}) + P_{\ominus}(A_t | Y_{0:t})}_{P(A_t | Y_{0:t})} \right) \quad (52)$$

578 From equation 52 to 53 we used the fact that independent regions of the marginal distributions will remain
 579 unchanged after an observation, $P_{\ominus}(A_t | Y_{0:t-1}) = P_{\ominus}(A_t | Y_{0:t})$, and before re-normalisation. This results
 580 in the final recursive update:

$$P(A_t, |Y_{0:t}) = P(A_t | Y_{0:t-1}) - \left(P_{\cap}(A_t | Y_{0:t-1}) - P_{\cap}(A_t | Y_{0:t}) \right) \quad (53)$$

581 Equation 53 states that only elements of the marginals which are dependent will change by the difference
 582 before and after a measurement update.

583 **8.6 Figures**

584 Frontiers requires figures to be submitted individually, in the same order as they are referred to in the
 585 manuscript. Figures will then be automatically embedded at the bottom of the submitted manuscript. Kindly
 586 ensure that each table and figure is mentioned in the text and in numerical order. Figures must be of
 587 sufficient resolution for publication see here for examples and minimum requirements. Figures which are
 588 not according to the guidelines will cause substantial delay during the production process. Please see here
 589 for full Figure guidelines

590 **8.7 Tables**

591 Tables should be inserted at the end of the manuscript. Please build your table directly in LaTeX. Tables
 592 provided as jpeg/tiff files will not be accepted. Please note that very large tables (covering several pages)
 593 cannot be included in the final PDF for reasons of space. These tables will be published as Supplementary
 594 Material on the online article page at the time of acceptance. The author will be notified during the
 595 typesetting of the final article if this is the case.

SUPPLEMENTAL DATA

REFERENCES

- 596 Arulampalam, M., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online
597 nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on* 50, 174–188.
598 doi:10.1109/78.978374
- 599 Bake, C., Tenenbaum, J., and Saxe, R. (2011). Bayesian theory of mind: Modeling joint belief-desire
600 attribution. *Journal of Cognitive Science*
- 601 Caralone, L., Du, J., Ng, M., Bona, B., and Indri, M. (2010). An application of kullback-leibler divergence
602 to active slam and exploration with particle filters. In *Intelligent Robots and Systems (IROS), 2010*
603 *IEEE/RSJ International Conference on*. 287–293. doi:10.1109/IROS.2010.5652164
- 604 Carrillo, H., Reid, I., and Castellanos, J. (2012). On the comparison of uncertainty criteria for active
605 slam. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. 2080–2087.
606 doi:10.1109/ICRA.2012.6224890
- 607 de Chambrier, G. and Billard, A. (2013). Learning search behaviour from humans. In *Robotics and*
608 *Biomimetics (ROBIO), 2013 IEEE International Conference on*. 573–580. doi:10.1109/ROBIO.2013.
609 6739521
- 610 Durrant-Whyte, H. and Bailey, T. (2006). Simultaneous localization and mapping: part i. *Robotics*
611 *Automation Magazine, IEEE* 13, 99–110. doi:10.1109/MRA.2006.1638022
- 612 González-Baños, H. H. and Latombe, J. (2002). Navigation strategies for exploring indoor environments. *I.*
613 *J. Robotic Res.* 21, 829–848. doi:10.1177/0278364902021010834
- 614 Grisetti, G., Kummerle, R., Stachniss, C., and Burgard, W. (2010). A tutorial on graph-based slam.
615 *Intelligent Transportation Systems Magazine, IEEE* 2, 31–43. doi:10.1109/MITS.2010.939925
- 616 Hoffman, J., Spranger, M., Gohring, D., and Jungel, M. (2005). Making use of what you don't see:
617 negative information in markov localization. In *Intelligent Robots and Systems, 2005. (IROS 2005).*
618 *2005 IEEE/RSJ International Conference on*. 2947–2952. doi:10.1109/IROS.2005.1545087
- 619 Hoffmann, J., Spranger, M., Gohring, D., Jungel, M., and Burkhard, H.-D. (2006). Further studies on the
620 use of negative information in mobile robot localization. In *Robotics and Automation, 2006. ICRA 2006.*
621 *Proceedings 2006 IEEE International Conference on*. 62–67. doi:10.1109/ROBOT.2006.1641162
- 622 Huang, Y. and Gupta, K. (2008). Rrt-slam for motion planning with motion and map uncertainty for robot
623 exploration. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference*
624 *on*. 1077–1082. doi:10.1109/IROS.2008.4651183
- 625 Kavraki, L., Svestka, P., Latombe, J.-C., and Overmars, M. (1996). Probabilistic roadmaps for path
626 planning in high-dimensional configuration spaces. *Robotics and Automation, IEEE Transactions on* 12,
627 566–580. doi:10.1109/70.508439
- 628 Kollar, T. and Roy, N. (2008). Efficient optimization of information-theoretic exploration in slam. In
629 *Proceedings of the 23rd National Conference on Artificial Intelligence - Volume 3* (AAAI Press),
630 AAAI'08, 1369–1375
- 631 Lidoris, G. (2011). *State Estimation, Planning, and Behavior Selection Under Uncertainty for Autonomous*
632 *Robotic Exploration in Dynamic Environments* (Kassel University: kassel university press GmbH)
- 633 Montemerlo, M. and Thrun, S. (2003). Simultaneous localization and mapping with unknown data
634 association using fastslam. In *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE*
635 *International Conference on*. vol. 2, 1985–1991 vol.2. doi:10.1109/ROBOT.2003.1241885
- 636 Montemerlo, M., Thrun, S., Koller, D., and Wegbreit, B. (2003). Fastslam 2.0: An improved particle
637 filtering algorithm for simultaneous localization and mapping that provably converges. In *In Proc. of the*
638 *Int. Conf. on Artificial Intelligence (IJCAI)*. 1151–1156

- 639 Plagemann, C., Kersting, K., Pfaff, P., and Burgard, W. (2007). Gaussian beam processes: A nonparametric
640 bayesian measurement model for range finders. In *In Proc. of Robotics: Science and Systems (RSS)*
641 Ross, S., Pineau, J., Paquet, S., and Chaib-draa, B. (2008). Online planning algorithms for pomdps. *Journal*
642 *Artificial Intelligence Research* 32, 663–704
- 643 Shachter, R. D. (1998). Bayes-ball: Rational pastime (for determining irrelevance and requisite information
644 in belief networks and influence diagrams). In *Proceedings of the Fourteenth Conference on Uncertainty*
645 *in Artificial Intelligence* (San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.), UAI'98,
646 480–487
- 647 Stachniss, C., Grisetti, G., and Burgard, W. (2005). Information gain-based exploration using rao-
648 blackwellized particle filters. In *Proc. of Robotics: Science and Systems (RSS)* (Cambridge, MA,
649 USA)
- 650 Thrun, S. (2002). Particle filters in robotics. In *in Proceedings of the 17th Annual Conference on*
651 *Uncertainty in AI (UAI)*
- 652 Thrun, S. and Bü, A. (1996). Integrating grid-based and topological maps for mobile robot navigation. In
653 *Proceedings of the Thirteenth National Conference on Artificial Intelligence - Volume 2* (AAAI Press),
654 AAAI'96, 944–950
- 655 Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic Robotics (Intelligent Robotics and Autonomous*
656 *Agents)* (Cambridge, MA: The MIT Press)
- 657 Thrun, S. and Leonard, J. J. (2008). Simultaneous localization and mapping. In *Springer Handbook of*
658 *Robotics*. 871–889. doi:http://dx.doi.org/10.1007/978-3-540-30301-5_38
- 659 Valencia, R., Miro, J., Dissanayake, G., and Andrade-Cetto, J. (2012). Active pose slam. In *Intelligent*
660 *Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. 1885–1891. doi:[10.1109/IROS.2012.6385637](https://doi.org/10.1109/IROS.2012.6385637)

FIGURES