Control of noisy differential-drive vehicles from time-bounded temporal logic specifications

This work was made openly accessible by BU Faculty. Please share how this access benefits you.
Your story matters.

| Version | Citation (published version): |
|---------|-------------------------------|
|         | Igor Cizelj, Calin Belta. 2013. "Control of Noisy Differential-Drive Vehicles from Time-Bounded Temporal Logic Specifications." 2013 IEEE International Conference on Robotics and Automation (ICRA), pp. 2021 - 2026 (6). |

https://hdl.handle.net/2144/29967

Boston University
Control of Noisy Differential-Drive Vehicles from Time-Bounded Temporal Logic Specifications

Igor Cizelj and Calin Belta

Abstract—We address the problem of controlling a noisy differential drive mobile robot such that the probability of satisfying a specification given as a Bounded Linear Temporal Logic (BLTL) formula over a set of properties at the regions in the environment is maximized. We assume that the vehicle can determine its precise initial position in a known map of the environment. However, inspired by practical limitations, we assume that the vehicle is equipped with noisy actuators and, during its motion in the environment, it can only measure the angular velocity of its wheels using limited accuracy incremental encoders. Assuming the duration of the motion is finite, we map the measurements to a Markov Decision Process (MDP). We use recent results in Statistical Model Checking (SMC) to obtain an MDP control policy that maximizes the probability of satisfaction. We translate this policy to a vehicle feedback control strategy and show that the probability that the vehicle satisfies the specification in the environment is bounded from below by the probability of satisfying the specification on the MDP. We illustrate our method with simulations and experimental results.

I. INTRODUCTION

Robot motion planning and control has been widely studied in the last twenty years. In “classical” motion planning problems [LaV06], the specifications are usually restricted to simple primitives of the type “go from A to B and avoid obstacles”, where A and B are two regions of interest in some environment. Recently, temporal logics, such as Linear Temporal Logic (LTL) and Computational Tree Logic (CTL) ([BK08], [CGP99]) have become increasingly popular for specifying robotic tasks (see, for example [LK04], [KGF07], [KF08], [KB08b], [WTM09], [BMKV11]). It has been shown that temporal logics can serve as rich languages capable of specifying complex motion missions such as “go to region A and avoid region B unless regions C or D are visited”.

In order to use existing model checking tools for motion planning (see [BK08]), many of the above-mentioned works rely on the assumption that the motion of the vehicle in the environment can be modeled as a finite system [CGP99] that is either deterministic (applying an available action triggers a unique transition [DLB12]) or nondeterministic (applying an available action can enable multiple transitions, with no information on their likelihoods [KB08a]). Recent results show that, if sensor and actuator noise models can be obtained from empirical measurements or an accurate simulator, then the robot motion can be modeled as a Markov Decision Process (MDP), and probabilistic temporal logics, such as Probabilistic CTL (PCTL) and Probabilistic LTL (PLTL), can be used for motion planning and control (see [LAB12]).

However, robot dynamics are normally described by control systems with state and control variables evaluated over infinite domains. A widely used approach for temporal logic verification and control of such a system is through the construction of a finite abstraction ([TP06], [Gir07], [KB08b], [YTC+12]). Even though recent works discuss the construction of abstractions for stochastic systems [JP09], [ADBS08], [DABS08], the existing methods are either not applicable to robot dynamics or are computationally infeasible given the size of the problem in most robotic applications.

In this paper, we consider a vehicle whose performance is measured by the completion of time constrained temporal logic tasks. In particular, we provide a conservative solution to the problem of controlling a stochastic differential drive mobile robot such that the probability of satisfying a specification given as a Bounded Linear Temporal Logic (BLTL) formula over a set of properties at the regions in the environment is maximized. Motivated by a realistic scenario of an indoor vehicle leaving its charging station, we assume that the vehicle can determine its precise initial position in a known map of the environment. The actuator noise is modeled as a random variable with a continuous probability distribution supported on a bounded interval, where the distribution is obtained through experimental trials. Also, we assume that the vehicle is equipped with two limited accuracy incremental encoders, each measuring the angular velocity of one of the wheels, as the only means of measurement available.

Assuming the duration of the motion is finite, through discretization, we map the incremental encoder measurements to an MDP. By relating the MDP to the vehicle motion in the environment, the vehicle control problem becomes equivalent to the problem of finding a control policy for an MDP such that the probability of satisfying the BLTL formula is maximized. Due to the size of the MDP, finding the exact solution is prohibitively expensive. We trade-off correctness for scalability and we use computationally efficient techniques based on sampling. Specifically, we use recent results in Statistical Model Checking for MDPs ([HMZ+12]) to obtain an MDP control policy and a Bayesian Interval Estimation (BIE) algorithm ([ZPC10]) to estimate the probability of satisfying the specification. We show that the probability that the vehicle satisfies the specification
in the original environment is bounded from below by the maximum probability of satisfying the specification on the MDP under the obtained control policy.

The main contribution of this work lies in bridging the gap between low level sensory inputs and high level temporal logic specifications. We develop a framework for the synthesis of a vehicle feedback control strategy from such specifications based on a realistic model of an incremental encoder. This paper extends our previous work ([CB12]) of controlling a stochastic version of Dubins vehicle such that the probability of satisfying a temporal logic statement, given as a PCTL formula, over some environmental properties, is maximized. Specifically, the approach presented here allows for richer temporal logic specifications, where the vehicle performance is measured by the completion of time constrained temporal logic tasks. Additionally, in order to deal with the increase in the size of the problem we use computationally efficient techniques based on sampling. In [HMZ+12], the authors use Statistical Model Checking for MDPs to solve a motion planning problem for a vehicle moving on a finite grid and knowing its state precisely, at all times, when the task is given as a BLTL formula. We adopt this approach to control a vehicle with continuous dynamics and allowing for uncertainty in its state.

The remainder of the paper is organized as follows. In Sec. II we introduce the necessary notation and review some preliminaries. We formulate the problem and outline the approach in Sec. III. In Sec. IV - VII we explain the construction of the MDP and the relation between the MDP and the motion of the vehicle in the environment. The vehicle control policy is obtained in Sec. VIII. Case studies and experimental results illustrating our approach are presented in Sec. IX. We conclude with final remarks and directions for future work in Sec. X.

II. PRELIMINARIES

In this section we provide a short and informal introduction to Markov Decision Processes (MDP) and Bounded Linear Temporal Logic (BLTL). For details about MDPs the reader is referred to [BK08] and for more information about BLTL to [JCL+09] and [ZPC10].

**Definition 1 (MDP):** A Markov Decision Process (MDP) is a tuple $M = (S, s_0, Act, A, P)$, where $S$ is a finite set of states; $s_0 \in S$ is the initial state; $Act$ is a finite set of actions; $A : S \rightarrow 2^{Act}$ is a function specifying the enabled actions at a state; $P : S \times Act \times S \rightarrow [0, 1]$ is a transition probability function such that for all states $s \in S$ and actions $a \in A(s)$: $\sum_{s' \in S} P(s, a, s') = 1$, and for all actions $a \notin A(s)$ and $s' \in S$: $P(s, a, s') = 0$.

A control policy for an MDP resolves nondeterminism in each state $s$ by providing a distribution over the set of actions enabled in $s$.

**Definition 2 (MDP Control Policy):** A control policy $\mu$ of an MDP $M$ is a function $\mu(s, a) : S \times Act \rightarrow [0, 1]$, s.t., $\sum_{a \in Act} \mu(s, a) = 1$ and $\mu(s, a) > 0$ only if $a$ is enabled in $s$. A control policy for which either $\mu(s, a) = 1$ or $\mu(s, a) = 0$ for all pairs $(s, a) \in S \times Act$ is called deterministic.

We employ Bounded Linear Temporal Logic (BLTL) to describe high level motion specifications. BLTL is a variant of Linear Temporal Logic (LTL) ([BK08]) which requires only paths of bounded size. A detailed description of the syntax and semantics of BLTL is beyond the scope of this paper and can be found in [JCL+09] and [ZPC10]. Roughly, formulas of BLTL are constructed by connecting properties from a set of proposition $\Pi$ using Boolean operators ($\neg$ (negation), $\land$ (conjunction), $\lor$ (disjunction)), and temporal operators ($U$ (bounded until), $F$ (bounded finally), and $G$ (bounded globally), where $r \in \mathbb{R}^{\geq 0}$ is the time bound parameter). The semantics of BLTL formulas are given over infinite traces $\sigma = (o_1, t_1)(o_2, t_2) \ldots$, $o_i \in 2^O$, $t_i \in \mathbb{R}^{\geq 0}$, $i \geq 1$, where $o_i$ is the set of satisfied propositions and $t_i$ is the time spent satisfying $o_i$. A trace satisfies a BLTL formula $\phi$ if $\phi$ is true at the first position of the trace; $F\phi$ means that $\phi$ will be true within $t$ time units; $G\phi$ means that $\phi$ will remain true for the next $t$ time units; and $\phi U \phi$ means that $\phi$ will be true within the next $t$ time units and $\phi$ remains true until then. More expressivity can be achieved by combining the above temporal and Boolean operators.

III. PROBLEM FORMULATION AND APPROACH

A. Problem Formulation

A differential drive mobile robot ([LaV06]) is a vehicle having two main wheels, each of which is attached to its own motor, and a third wheel which passively rolls along preventing the robot from falling over. In this paper, we consider a stochastic version of a differential drive mobile robot, which captures actuator noise:

$$
\begin{align*}
\dot{x} &= \frac{1}{L}(u_r + \epsilon_r + u_l + \epsilon_l) \cos(\theta) \\
\dot{y} &= \frac{1}{L}(u_r + \epsilon_r + u_l + \epsilon_l) \sin(\theta) \\
\dot{\theta} &= \frac{1}{L}(u_r - u_l - \epsilon_l),
\end{align*}
$$

where $(x, y) \in \mathbb{R}^2$ and $\theta \in [0, 2\pi]$ are the position and orientation of the vehicle in a world frame, $u_r$ and $u_l$ are the control inputs (angular velocities before being corrupted by noise), $U_r$ and $U_l$ are control constraint sets, and $\epsilon_r$ and $\epsilon_l$ are random variables modeling the actuator noise with continuous probability density functions supported on the bounded intervals $[\epsilon_r^{\min}, \epsilon_r^{\max}]$ and $[\epsilon_l^{\min}, \epsilon_l^{\max}]$, respectively. $L$ is the distance between the two wheels and $r$ is the wheel radius. We denote the state of the system by $q = [x, y, \theta]^T \in SE(2)$.

Motivated by the fact that the time optimal trajectories for the bounded velocity differential drive robots are composed only of turns in place and straight lines ([BM00]), we assume $U_r$ and $U_l$ are finite, but we make no assumptions on the optimality. We define

$$W_i = \{u + \epsilon | u \in U_i, \epsilon \in [\epsilon_i^{\min}, \epsilon_i^{\max}]\}, \quad i \in \{r, l\},$$

as the sets of applied control inputs, i.e., the sets of angular wheel velocities that are applied to the system in the presence of noise. We assume that time is uniformly discretized (partitioned) into stages (intervals) of length $\Delta t$, where stage $k$ is from $(k-1)\Delta t$ to $k\Delta t$. The duration of the motion is
finite and it is denoted $K\Delta t$ (later in this section we explain how $K$ is determined). We denote the control inputs and the applied control inputs at stage $k$ as $u_i^k \in U_i$, $i \in \{r,l\}$, and $\bar{w}_i^k \in W_i$, $i \in \{r,l\}$, respectively.

We assume that the vehicle is equipped with two incremental encoders, each measuring the applied control input (i.e., the angular velocity corrupted by noise) of one of the wheels. Motivated by the fact that the angular velocity is considered constant inside the given observation stage ([PTPZ07]), the applied controls are considered piecewise constant, i.e., $w_i^k : [\Pi_k \Delta t, k\Delta t] \rightarrow W_i$, $i \in \{r,l\}$, are constant over each stage.

Incremental encoder model: As shown in [PTPZ07], the measurement resolution of an incremental encoder is constant and for encoder $i$ we denote it as $\Delta \varepsilon_i$, $i \in \{r,l\}$. Given $\Delta \varepsilon_i$ and $[\varepsilon_i^{min}, \varepsilon_i^{max}]$, $i \in \{r,l\}$, then the following holds: $\exists n_i \in \mathbb{Z}^+$ s.t. $n_i \Delta \varepsilon_i = |\varepsilon_i^{max} - \varepsilon_i^{min}|$, $i \in \{r,l\}$. For more details see Sec. IX where we also explain how to obtain the measurement resolutions and the probability density functions. Then, $[\varepsilon_i^{min}, \varepsilon_i^{max}]$ can be partitioned into $n_i$ noise intervals of length $\Delta \varepsilon_i = [\varepsilon_i^j, \varepsilon_i^{j+1}]$, $j = 1, ..., n_i$, $i \in \{r,l\}$. We denote the set of all noise intervals $\delta_i = \{[\varepsilon_i^1, \varepsilon_i^{j+1}]\}$, $i \in \{r,l\}$. At stage $k$, if the applied control input is $u_i^k + \varepsilon_i$, the incremental encoder $i$ will return measured interval

$$[w_i^k, \bar{w}_i^k] = [u_i^k + \varepsilon_i, u_i^k + \bar{\varepsilon}_i],$$

where $\varepsilon_i \in [\varepsilon_i^j, \bar{\varepsilon}_i^j], i \in \{r,l\}$. In Fig. 1 we give an example. The pair of measured intervals at stage $k$, $([w_i^k, \bar{w}_i^k], [w_i^k, \bar{w}_i^k])$, returned by the incremental encoders, is denoted $\mathbb{W}^k$.

![Fig. 1](image)

The vehicle moves in a planar environment in which a set of non-overlapping regions of interest, denoted $R$, is present. Let $\Pi$ be the set of propositions satisfied at the regions in the environment. One of these propositions, denoted by $\pi_a \in \Pi$, signifies that the corresponding regions are unsafe. In this work, the motion specification is expressed as a BLTL formula $\phi$ over $\Pi$:

$$\phi = -\pi_a U^{T_1} (\phi_1 \land -\pi_a U^{T_1} (\phi_2 \land \ldots \land -\pi_a U^{T_1} \phi_f)), (2)$$

1Throughout the paper, we relax the notion of partition by allowing the endpoints of the intervals to overlap.


\((o_1, t_1)(o_2, t_2)\ldots(o_i, t_i), o_i \in \Pi \cup \emptyset, t_i \in \mathbb{R}^{\geq 0}, i \geq 1\), where \(o_i\) is the satisfied proposition\(^2\) and \(t_i\) is the time spent satisfying \(o_i\), as time evolves. A trajectory \(q(t)\) satisfies BLTL formula \(\phi\) if and only if the generated trace satisfies the formula. Given \(\phi\), for the duration of the motion we use the smallest \(K \in \mathbb{Z}^+\) for which model checking a trace is well defined, i.e., the smallest \(K\) for which the maximum nested sum of time bounds (see [ZPC10]) is at most \(K\Delta t\).

\[ t_i = K\Delta t - \sum_{j=1}^{i-1} t_j, \quad \text{with} \quad t_0 = 0. \]

- Let for \(K\Delta t\), \(o_i\) be the current satisfied propositions. Then, \(t_i = K\Delta t - \sum_{j=1}^{i-1} t_j\).

A trajectory \(q(t)\) satisfies BLTL formula \(\phi\) (Eqn. (2)) if and only if the trace generated according to the rules stated above satisfies the formula. Note that, since the duration of the motion is finite, the generated trace is also finite. In [ZPC10] the authors show that BLTL requires only traces of bounded lengths. The fact that the trace \(\sigma\) satisfies \(\phi\) is denoted \(\sigma \models \phi\). Given a trace \(\sigma\), the \(i\)-th state of \(\sigma\), denoted \(\sigma_i\), is \((o_i, t_i), i = 1, \ldots, l\). We denote \(\sigma_{ij}\) as the finite subsequence of \(\sigma\) that starts in \(\sigma_i\). Finally, given a formula \(\phi\), we denote subformula \(\neg \pi_a U^l \phi_j\) as \(\phi_j, j = 1, \ldots, f\). Using the BLTL semantics one can derive the following conditions to determine whether \(\sigma \models \phi\):

**Definition 4 (Satisfaction conditions):** Given a trace \(\sigma\) and a BLTL formula \(\phi\) (Eqn. (2)), let for \(j \in \{1, \ldots, f\}\), \(i_j, k_j \in \mathbb{N}\) be such that for some \(n \in \{1, \ldots, n_j\}\) the following holds:

1. \(o_{i_j} + k_j \in \Gamma_{j}^n\),
2. for each \(i_j \leq i < i_j + k_j, o_i \neq \pi_a, \)
3. \(\sum_{j=i}^{i_j+k_j-1} t_j \leq T_j,\) and
4. \(t_{i_j + k_j} \geq \tau_j\).

Then, \(\sigma_{ij} \models \phi_j\). If \(\forall j \in \{1, \ldots, f\}, \exists i_j, k_j \in \mathbb{N}\) s.t. \(\sigma_{ij} \models \phi_j\) where \(i_j = i_1 + k_1\) with \(i_1 = 1\), then \(\sigma \models \phi\).

**Example 2:** Consider the environment and the sample state (position) trajectory shown in Fig. 2. Let \(\phi\) be as in Eq. (3) with the following numerical values for the time bounds: \(T_1 = 6.2, T_2 = 2.3, T_2 = 0.2, \) and \(T_3 = 2.3\). The trajectory generates trace \(\sigma = (0.6, 1.2)(\pi_0, 0.75)(0.0, 0.44)(\pi_5, 0.61)(0, 1.66)(\pi_4, 1.22)\).

The following holds: \(\sigma_{i} \models \phi_1\) since for \(i = 1 = k_1 = 1, o_2 \in \{\pi_0\}, o_1 \neq \pi_a, t_1 \leq T_1; \sigma_2 \models \phi_2\) since for \(t_2 = 2, k_2 = 2, o_2 \in \{\pi_5\}, o_2, o_3 \neq \pi_a, t_3 \leq T_2, t_2 \leq \tau_2; \) and \(\sigma_{4} \models \phi_4\) since for \(i_4 = 4 \) and \(k_3 = 2, o_6 \in \{\pi_4\}, o_4, o_5 \neq \pi_a\) and \(t_4 + t_5 \leq T_3\); Thus, \(\sigma \models \phi\).

### IV. Generating a Trace

In this section we explain how, given a state trajectory the corresponding trace is generated. Let us denote \([\pi] = \{(x, y) \in \mathbb{R}^2 | (x, y) \in \cup_{r \in R} r\}\) as the set of positions that satisfy proposition \(\pi\), where \(R \subseteq R\) is the set of regions labeled with proposition \(\pi\).

**Definition 3 (Generating a trace):** The trajectory corresponding to a state trajectory \(q(t) = [x(t), y(t), \theta(t)]^T\) is a finite sequence \(\sigma = (o_1, t_1)(o_2, t_2)\ldots(o_i, t_i), o_i \in \Pi \cup \emptyset, t_i \in [0, K\Delta t], i = 1, \ldots, l, l \geq 1\), where \(o_i\) is the satisfied proposition and \(t_i\) is the time spent satisfying \(o_i\), generated according to the following rules, for all \(t, t' \in [0, K\Delta t]\):

- \(o_1 = \pi \in \Pi\) if \((x(0), y(0)) \in [\pi]\) and \(o_1 = \emptyset\) otherwise.

- Let \(o_1\) be the satisfied proposition at some \(t\). Then:
  1. \(\text{If } o_1 = \emptyset, \text{ then } o_{i+1} = \pi \in \Pi, \text{ if} (i) \exists t' > t \text{ s.t. } (x(t'), y(t')) \in [\pi], \text{ and } (ii) \exists t \in [t, t') \text{ s.t. } (x(t), y(t)) \in [\pi], \forall t' \in [0, K\Delta t] \text{ and } t_i = \min_{t \in [t_i, K\Delta t]} \{t | (x(t), y(t)) \notin [\pi]\} - \sum_{j=0}^{i-1} t_j, \text{ with } t_0 = 0.\)
  2. \(\text{If } o_1 = \pi \in \Pi, \text{ then } o_{i+1} = \emptyset \text{ if } \exists t' > t \text{ s.t. } (x(t'), y(t')) \notin [\pi], \text{ and } t_i = \min_{t \in [t_i, K\Delta t]} \{t | (x(t), y(t)) \notin [\pi]\} - \sum_{j=0}^{i-1} t_j, \text{ with } t_0 = 0.\)

\(^2\)Since the regions of interest are non-overlapping it follows that \(o_i \in \Pi \cup \emptyset\).

### V. Construction of an MDP Model

Recall that \(\varepsilon_i\) is a random variable with a continuous probability density function supported on the bounded interval \([\varepsilon_i^{\min}, \varepsilon_i^{\max}]\), \(i \in \{r, l\}\). The probability density functions are obtained through experimental trials (see Sec. IX) and they are defined as follows:

\[
Pr(\varepsilon_i \in [\varepsilon_i^l, \varepsilon_i^u]) = \gamma_i^l, \quad (4) \quad \text{or} \quad Pr(\varepsilon_i \in [\varepsilon_i^l, \varepsilon_i^u]) = \gamma_i^u, \quad (4)
\]

An MDP \(M\) that captures every sequence realization of pairs of measurements returned by the incremental encoders is defined as a tuple \((S, \alpha, \text{Act}, A, P)\), where:

- \(S = \cup_{k=1}^{K} \{[u_r + \varepsilon_r, u_r + \varepsilon_l], [u_l + \varepsilon_l, u_l + \varepsilon_l] \} | u_r \in U_r, u_l \in U_l, [\varepsilon_r, \varepsilon_l] \in \varepsilon_i^l, [\varepsilon_l, \varepsilon_l] \in \varepsilon_i^l\}^K\). The meaning of the state is as follows: \((\gamma_i^l, \gamma_i^u) \in S\), means that at stage \(i\), \(1 \leq i \leq k\), the pair of measured intervals is \(\gamma_i^l, \gamma_i^u\).
- \(s_0 = \emptyset\) is the initial state.
• $Act = \{U_i \times U_j\} \cup \emptyset$ is the set of actions, where $\emptyset$ is a dummy action.

• $A : S \to 2^{Act}$ gives the enabled actions at state $s$: if $|s| = K$, i.e., if the termination time is reached, $A(s) = \emptyset$, otherwise $A(s) = \{U_i \times U_j\}$.

• $P : S \times Act \times S \to [0, 1]$ is a transition probability function constructed by the following rules:

  1) If $s = (\{y_i\}, \ldots, \{y_k\}) \in S$ then $P(s,a,s') = p_{s,1}^a p_{s,2}^a$ iff $s' = (\{y_i\}, \ldots, \{y_k\})$, $[u_i + \epsilon_i^a, u_i + \epsilon_i^a]$; $s$ and $a = (u_i, u_j) \in \{U_i \times U_j\}$ where $m = 1, \ldots, n, n = 1, \ldots, n, k = 1, \ldots, K$;

  2) If $|s| = K$ then $P(s,a,s') = 1$ iff $a = \emptyset$ and $s' = s$;

  3) $P(s,a,s') = 0$ otherwise.

Rule 1) defined above follows from the fact that given $u^k_i$ and $u^k_j$ as the control inputs at state $k$, the pair of measured intervals at state $k + 1$ is $([u_i + \epsilon_i^a, u_i + \epsilon_i^a], [u_j + \epsilon_j^a, u_j + \epsilon_j^a])$ with probability $p_{s,1}^a p_{s,2}^a$, since $Pr(e_i \in \{y_i\}, \{y_i\}) = p_{s,1}^a$ and $Pr(e_i \in \{y_i\}, \{y_i\}) = p_{s,2}^a$, which follows from Eqn. (4) (see the MDP fragment in Fig. 3). Rule 2) states that if the length of $s$ is equal to $K$, i.e., if the termination time is reached, then $A(s) = \emptyset$ with $P(s,a,s) = 1$.

![Fig. 3. A fragment of the MDP $M$ where $n_i = n_j = 2$. Thus, $p_{s,1}^a = Pr(e_i \in \{y_i\}, \{y_i\})$, for $m = 1, 2$, and $p_{s,2}^a = Pr(e_i \in \{y_i\}, \{y_i\})$, for $m = 1, 2$. Action $(u_i, u_j) \in A(s)$ enables four transitions. For example, given state $s = (\{y_i\})$, the new state is $\{y_i\} \times \{y_i\}$, where $W_2 = ([u_i - \epsilon_i^a, u_i + \epsilon_i^a], [u_i + \epsilon_i^a, u_i + \epsilon_i^a])$, with probability $p_{s,1}^a p_{s,2}^a$. This corresponds to applied control inputs being equal to $u_i + \epsilon_i$ and $u_i + \epsilon_i$ where $e_i \in \{y_i\}, \{y_i\}$.

**Proposition 1:** The model $M$ defined above is a valid MDP, i.e., it satisfies the Markov property and $P$ is a transition probability function.

**Proof:** The proof follows from construction of $P$. Given current state $s \in S$ and an action $a \in A(s)$, the conditional probability distribution of future states depends only on the current state $s$, not on the sequences of events that preceded it (see rule 1) above). Thus, the Markov property holds. In addition, since for every state and $s \in A(s)$: $\sum_{s' \in S} P(s,a,s') = \sum_{m=1}^{n+1} p_{s,1}^a p_{s,2}^a = \sum_{m=1}^{n+1} p_{s,1}^a p_{s,2}^a = 1$, it follows that $P$ is a valid transition probability function. $\blacksquare$

**VI. POSITION UNCERTAINTY**

**A. Nominal state trajectory**

For each interval belonging to the set of noise intervals $\mathcal{D}$, we define a representative value $\vec{e}_i = (\vec{e}_i^1 + \vec{e}_i^2)/2$, $j_i = 1, \ldots, n_i$, $i \in \{r, l\}$, i.e., $\vec{e}_i$ is the midpoint of interval $[\vec{e}_i^1, \vec{e}_i^2] \in \mathcal{D}$, $i \in \{r, l\}$. We denote the set of representative values as $E_i = \{\vec{e}_i, \ldots, \vec{e}_i^m\}, i \in \{r, l\}$.

We use $q^k(t)$, $d^k_i$, and $w^k_i$, $t \in [k-1] \Delta, k \Delta]$, $k = 1, \ldots, K$, to denote the state trajectory and the constant applied controls at stage $k$, respectively. With a slight abuse of notation, we use $q^k$ to denote the end of state trajectory $q^k(t)$, i.e., $q^k = q^k(k \Delta)$. Given state $q^k - 1$, the state trajectory $q^k(t)$ can be derived by integrating the system given by Eqn. (1) from the initial state $q^k - 1$, and taking into account the applied controls are constant and equal to $w^k_i$ and $w^k_i$. Throughout the paper, we will also denote this trajectory by $q^k(q^k - 1, w^k_i, w^k_i, t)$, when we want to explicitly capture the initial state $q^k - 1$ and the constant applied controls $w^k_i$ and $w^k_i$.

Given a path through the MDP:

$$s_0(u_i,u_j) \xrightarrow{s_1(u_i^2,u_j^2)} \cdots \xrightarrow{s_{K-1}(u_i^k,u_j^k)} s_K,$$

(5)

where $s_k = (\{y_i\}, \ldots, \{y_k\}$, with $\{y_k\} = ([u_i + \epsilon_i^a, u_i + \epsilon_i^a], [u_j + \epsilon_j^a, u_j + \epsilon_j^a]), k = 1, \ldots, K$, we define the nominal state trajectory $q(t)$, $t \in [0, K \Delta]$, as follows:

$$q(t) = q^k(q^k - 1, u_i, u_j, e_i^k, e_j^k, t), t \in [k-1] \Delta, k \Delta],$$

where $e_i^k \in E_i$ is such that $e_i^k \in \{e_i, \vec{e}_i\}$, $i \in \{r, l\}$, and $q_0 = q_{init}$. For every path through the MDP, its nominal state trajectory is well defined. The next step is to define the uncertainty evolution, along the nominal state trajectory, since the applied controls can take any value within the measured intervals.

**B. Position uncertainty evolution**

Since a motion specification is a statement about the propositions satisfied by the regions of interest in the environment, in order to answer whether some state trajectory satisfies BLTL formula $\phi$ it is sufficient to know its projection in $\mathbb{R}^2$. Therefore, we focus only on the position uncertainty.

The position uncertainty of the vehicle when its nominal position is $(x, y) \in \mathbb{R}^2$ is modeled as a disc centered at $(x, y)$ with radius $d \in \mathbb{R}$, where $d$ denotes the distance uncertainty:

$$D(x, y, d) = \{(x', y') \in \mathbb{R}^2||\|(x, y), (x', y')\|| \leq d\},$$

(6)

where $\|\cdot\|$ denotes the Euclidean distance. Next, we explain how to obtain $d$.

First, let $\Delta \theta \in S_1$ denote the orientation uncertainty. Let $q(t)$, $t \in [0, K \Delta]$, be the nominal state trajectory corresponding to a path through the MDP (Eqn. (5)). Then, $q(t)$ can be partitioned into $K$ state trajectories: $q^k(t) = q^k(q^k - 1, u_i, u_j, e_i^k, e_j^k, t), t \in [k-1] \Delta, k \Delta], k = 1, \ldots, K$, where $e_i^k \in E_i$ is such that $e_i^k \in \{e_i, \vec{e}_i\}, i \in \{r, l\}$, and $q_0 = q_{init}$ (see Fig. 4). The distance and orientation uncertainty at state $q^k$ are denoted as $d^k$ and $\Delta \theta^k$, respectively. We set $d^k$ and $\Delta \theta^k$ at state $q^k = \{x^k, y^k, \theta^k\}$ to equal:

$$d^k = \max_{(x', y', \theta') \in \mathbb{R}^2, (x', y', \theta') \in \mathbb{R}^2, (x', y', \theta')} \|\|(x', y', \theta')\|| - (x', y', \theta')\||,$$

$$\Delta \theta^k = \max_{(x', y', \theta') \in \mathbb{R}^2, (x', y', \theta') \in \mathbb{R}^2, (x', y', \theta')} \|\|(x', y', \theta')\|| - (x', y', \theta')\||,$$

(7)

where $\mathbb{R}^k = \{q^k([x^k - 1, y^k - 1, \theta^k - 1 + \alpha]^T, u_i, e_i, \theta_i, k \Delta)\}$

$\alpha \in \{\Delta \theta^k - 1, -\Delta \theta^k - 1\}, e_i \in \{e_i, \vec{e}_i\}$, $e_j \in \{e_j, \vec{e}_j\}$, $\rho$ for $k = 1, \ldots, K$, where $\rho = 0$ and $\Delta \theta^0 = 0$. 







Eqn. (4) and (5) are obtained using a worst scenario assumption. At stage \( k \), the pair of measured intervals is \( \mathcal{W} \left( \left[ u_i^k + \varepsilon_i^k, u_i^k + \varepsilon_i^k \right] \right) \) and we use the endpoints of the measured intervals to set \( \mathcal{R}^k \). \( \mathcal{R}^k \) is the smallest set of points in \( \mathcal{SE}(2) \), at the end of stage \( k \), guaranteed to contain (i) the state with the maximum distance (in Euclidean sense) from \( q^k \) given that the applied controls at stage \( i \) are within the measured intervals at stage \( i \), and (ii) the state with the maximum orientation difference compared to \( q^k \) given that the applied controls at stage \( i \) are within the measured intervals at stage \( i, i = 1, \ldots, k \). (For more details about \( \mathcal{R}^k \) see [FMA99]). An example is given in Fig. 4.

From Eqn. (7) and (8) it follows that, given a nominal state trajectory \( q(t) \), \( t \in [0,K \Delta t] \), the distance uncertainty increases as a function of time. The way it changes along \( q(t) \) makes it difficult to characterize the exact shape of the position uncertainty region. Instead, we use a conservative approximation of the region. We define \( d : [0,K \Delta t] \rightarrow R \) as an approximate distance uncertainty trajectory and we set \( d(t) = d^k, t \in [k-1] \Delta t, k \Delta t, k = 1, \ldots, K \), i.e., we set the distance uncertainty along the state trajectory \( q^k(t) \) equal to the maximum value of the distance uncertainty along \( q^k(t) \), which is at state \( q^k \). An example illustrating this idea is given in Fig. 4.

**Proposition 2:** Given a path through the MDP \( M \) (Eqn. (5)), and the corresponding \( q(t) \) and \( d(t) \), \( t \in [0,K \Delta t] \), as defined above, then any state trajectory \( q' \) of \( \sigma = q(t) \) satisfies \( d(t) \) equal to the maximum value of the distance uncertainty along \( q^k(t) \), which is at state \( q^k \). An example illustrating this idea is given in Fig. 4.

**VII. Generating a trace under the position uncertainty**

Let \( q(t) \) be a nominal state trajectory with the distance uncertainty trajectory \( d(t) \), \( t \in [0,K \Delta t] \). In this subsection we introduce a set of conservative rules according to which the trace corresponding to the uncertainty region \( \mathcal{D}(\gamma) ^2(\gamma) \) is generated. This rules guarantee that if the generated trace satisfies \( \phi \) (Eqn. (2)) then any state (position) trajectory, inside \( \mathcal{D}(\gamma) ^2(\gamma) \), will satisfy \( \phi \).

**Definition 5 (Generating a trace under uncertainty):**

The trace corresponding to an uncertainty region \( \mathcal{D}(\gamma) ^2(\gamma) \) is a finite sequence \( \sigma = (o_1, t_1) (o_2, t_2) \ldots (o_{l+1}, t_{l+1}) \), \( o_i \in \Pi \cup \emptyset, t_i \in [0,K \Delta t], i = 1, \ldots, l, l \geq 1 \), where \( o_i \) is the satisfied proposition and \( t_i \) is the time spent satisfying \( o_i \), generated according to the following rules, for all \( t, t', \tau \in [0,K \Delta t] \):

- **1)** If \( o_i = \pi \in \Pi \setminus \pi_o \), then \( o_{i+1} = \emptyset \) if \( \exists \tau > t \) s.t. \( D((x(t), y(t)), d(t)) \subseteq [\pi] \), \( t_i = \min_{\tau \in [0,K \Delta t]} \{ t | D((x(t), y(t)), d(t)) \subseteq [\pi] \} - \sum_{j=0}^{i-1} t_j \), with \( t_0 = 0 \).
- **2)** If \( o_i = \pi_o \), then \( o_{i+1} = \emptyset \) if \( \exists \tau' > t \) s.t. \( D((x(t'), y(t')), d(t')) \cap [\pi] = \emptyset \) and \( t_i = \min_{\tau \in [0,K \Delta t]} \{ t | D((x(t), y(t)), d(t)) \subseteq [\pi] \} - \sum_{j=0}^{i-1} t_j \), with \( t_0 = 0 \).
- **3)** If \( o_i = \emptyset \), then \( o_{i+1} = \pi \), if \( \exists \tau > t \) s.t. \( D((x(t'), y(t'), d(t')) \subseteq [\pi] \), with \( \forall \pi' \in \Pi \setminus \pi_o \).
- **4)** If \( o_i = \emptyset \), then \( o_{i+1} = \pi \), if \( \exists \tau > t \) s.t. \( D((x(t'), y(t'), d(t')) \subseteq [\pi] \), with \( \forall \pi' \in \Pi \setminus \pi_o \) and \( \forall \pi' \in \Pi \setminus \pi_o \).

- Let for \( K \Delta t \), \( o_i \) be the current satisfied proposition. Then \( t_i = K \Delta t - \sum_{j=0}^{i-1} t_j \), with \( t_0 = 0 \).

In Fig. 5 we show an uncertainty region and the corresponding trace generated according to rules stated above. Next, we show that if the trace corresponding to an uncertainty region satisfies \( \phi \), then any state (position) trajectory inside the uncertainty region also satisfies \( \phi \).

**Proposition 3:** Let \( D((x(t), y(t)), d(t)) \) be the uncertainty region corresponding to a path through the MDP \( M \) (Eqn. (5)) and let \( q' \) be any state trajectory as defined in Prop. 2. Let \( \sigma^D = (o_1^D, t_1^D) \ldots (o_{l^D}^D, t_{l^D}^D) \) and \( \sigma^q = (o_1^q, t_1^q) \ldots (o_{l^q}^q, t_{l^q}^q) \) be the corresponding traces. Given BLTL formula \( \phi \) (Eqn. (2)), if \( \sigma^D \models \phi \), then \( \sigma^q \models \phi \).
Proof: First, we state two relations between the given traces:

1) Let \( \sigma^0 = \pi \in \Pi \setminus \pi_a \) for some \( i \in \{1, \ldots, k\} \). Then, the following holds: \( \exists j \in \{1, \ldots, l\} \) such that \( \sigma^0_j = \pi \) and \( t^0_j \leq t^1_j \).

Informally, if \( t^0_j \) is the time \( D((x(t), y(t)), d(t)) \) spent inside the region satisfying proposition \( \pi \), then \( d'(t) \) will spend at least \( t^0_j \) time units inside that region.

2) Let \( \sigma^0 = \pi \in \Pi \setminus \pi_a \) and \( \sigma^0 = \pi' \in \Pi \setminus \pi_a \) for some \( i, j \in \{1, \ldots, l\} \). Then, the following holds: \( \exists j, j' \in \{1, \ldots, l\}, j' > j \) such that \( \sigma^0_j = \pi \) and \( \sigma^0_j = \pi' \).

In addition, \( \sum_{i=1}^{l} t^0_i \leq \sum_{i=1}^{l} t^1_i \).

Informally, if the time between \( D((x(t), y(t)), d(t)) \) entering a region satisfying \( \pi \) and then entering a region satisfying \( \pi' \) is \( \sum_{i=1}^{l} t^0_i \) time units, then the time between \( d'(t) \) entering the region satisfying \( \pi \) and then entering the region satisfying \( \pi' \) is bounded from above by \( \sum_{i=1}^{l} t^0_i \). For more intuition about these relations see Fig. 5.

VIII. VEHICLE CONTROL STRATEGY

Given the MDP \( M \), the next step is to obtain a control policy that maximizes the probability of generating a path through \( M \) such that the corresponding trace (as defined in Sec. VI and VII) is satisfying. There are existing approaches that, given an MDP and a temporal logic formula, generate an exact control policy that maximizes the probability of satisfying the specification. In general, exact techniques rely on reasoning about the entire state space, which is a limiting factor in their applicability to large problems. Given \( U_t, n_t, n_t \) and \( K \), the size of the MDP is bounded above by \( |U_t| \times |U_t| \times |n_t| \times n_t \). For even a single case study, due to the size of \( M \), using the exact methods to obtain a control policy is computationally too expensive. Therefore, we decide to trade-off correctness for scalability and use computationally efficient techniques based on system sampling.

A. Overview

We obtain a suboptimal control policy by iterating over the control synthesis and the probability estimation procedure until the stopping criterion is met (see Sec. VIII-C). In the control synthesis procedure we use the control synthesis approach from [HMZ+12] to generate a control policy for the MDP \( M \). In particular we use a control policy optimization part of the algorithm which consists of the control policy evaluation and the control policy improvement procedure to incrementally improve a candidate control policy (control policy is initialized with a uniform distribution at each state). Next, in the probability estimation procedure we use SMC by BIE, as presented in [ZPC10]. We estimate the probability that the MDP \( M \), under the candidate control policy, generates a path such that the corresponding trace satisfies BLTL formula \( \phi \). Finally, if the estimated probability converges, i.e., if the stopping criterion is met, we map the control policy to a vehicle control strategy. Otherwise, the control synthesis procedure is restarted using the latest update of the control policy. The flow of this approach is depicted in Fig. 6.

B. Control synthesis

The details of the control policy optimization algorithm can be found in [HMZ+12] and here we only give an informal overview of the approach. In the control policy evaluation procedure we sample paths of the MDP \( M \) under the current control policy \( \mu \). Given a path \( \omega = s_0 \xrightarrow{a_1} s_1 \xrightarrow{a_2} s_2 \ldots \xrightarrow{a_k} s_K \), where \( a_k = (u^k_i, u^k_j) \), the corresponding trace \( \sigma \) is generated as described in Sec. VI and VII. Next, we check formula \( \phi \) on each \( \sigma \) and estimate how likely it is for each action to lead to the satisfaction of BLTL formula \( \phi \), i.e., we obtain the estimate of the probability that a path crossing a state-action pair, \( (s^k, a^{k+1}) \), \( k = 0, \ldots, K - 1 \), in \( \omega \) will generate a trace that satisfies \( \phi \). These estimates are then used in the control policy improvement procedure, in which we update the control policy \( \mu \) by reinforcing the actions that led to the satisfaction of \( \phi \) most often.

Assuming \( \sigma^0 \models \phi \), then \( \forall j \in \{1, \ldots, f\} \), \( \exists i, j, k \in \mathbb{N} \) and some \( n \in \{1, \ldots, n_j\} \) such that \( \sigma^{f}_j \models \phi \) (see Def. 3). Then, from Prop. 2 and Def. 3 it follows that \( \forall j \in \{1, \ldots, f\} \), \( \exists i, j, k \in \mathbb{N} \) such that:

1) \( \sigma^{f}_{j+z} \in \Pi^n \),
2) for each \( s_j \leq i < s_j + z_j \), \( \sigma^q_i \neq \pi_a \),
3) \( \sum_{i=z}^{j+z} t^q_i \leq \sum_{i=j}^{j+z} t^p_i \leq t_j \) (2nd relation above),
4) \( t^q_{j+z} \geq t^p_{j+z} \) (1st relation above).

where \( s_j = s_j + z_j \) with \( s_j = 1 \).

Thus, \( \forall j \in \{1, \ldots, f\} \), \( \sigma^q_j \models \phi \), and according to Def. 3 it follows that \( \sigma^q \models \phi \). In Fig. 5 we give an example.
In words, we compute a control policy that always picks the best estimated action at each state.

**C. Probability estimation**

Next, we determine the estimate of the probability that the MDP $M$, under the deterministic control policy $\mu_{det}$, generates a path such that the corresponding trace satisfies BLTL formula $\phi$. To do so we use the BIE algorithm as presented in [ZPC10]. We denote the exact probability as $p_M$ and the estimate as $\hat{p}_M$.

The inputs of the algorithm are the MDP $M$, control policy $\mu_{det}$, BLTL formula $\phi$, half interval size $\delta \in (0, \frac{1}{2})$, interval coefficient $c \in (\frac{1}{2}, 1)$, and the coefficients $\alpha, \beta$ of the Beta prior. The algorithm returns $\hat{p}_M$. The algorithm generates traces by sampling paths through $M$ under $\mu_{det}$ (as described in Sec. VI and VII) and checks whether the corresponding traces satisfy $\phi$, until enough statistical evidence has been found to support the claim that $p_M$ is inside the interval $[\hat{p}_M - \delta, \hat{p}_M + \delta]$ with arbitrarily high probability, i.e., $\Pr(p_M \in [\hat{p}_M - \delta, \hat{p}_M + \delta]) \geq c$.

We stop iterating over the control synthesis and the probability estimation procedure when the difference between the two consecutive probability estimates converges to a neighborhood of radius $c \in (0, 1)$, i.e., when the difference is smaller or equal to $c$. Let $\mu_{det}$ and $\hat{p}_M$ be the current control policy and the corresponding probability estimate, respectively, when the stopping criterion is met.

**D. Control strategy**

The vehicle control strategy is a function $\gamma: S \to \{U_t \times U_t\}$ that maps a sequence of pairs of measured intervals, i.e., a state of the MDP, to the control inputs:

$$\gamma((\mathbb{W}^1, \ldots, \mathbb{W}^k)) = \gamma(s_k) = \max_{a \in \text{Act}(s_k)} \mu_{det}^*(s_k, a), \quad (9)$$

$k = 1, \ldots, K - 1$ with $\gamma(s_0) = \max_{a \in \text{Act}(s_0)} \mu_{det}^*(s_0, a)$.

At stage $k$, the control inputs are

$$(u_t^k, u_t^k) = \gamma((\mathbb{W}^1, \ldots, \mathbb{W}^{k-1})) \in \{U_t \times U_t\}.$$

Thus, given a sequence of pairs of measured intervals, $\gamma$ returns the control inputs for the next stage; the control inputs are equal to the action returned by $\mu_{det}$ at the state of the MDP corresponding to that sequence.

**Theorem 1**: The probability that the system given by Eqn. (1), under the vehicle control strategy $\gamma$, generates a state trajectory that satisfies BLTL formula $\phi$ (Eqn. (2)) is bounded from below by $p_M$, where $\Pr(p_M \in [\hat{p}_M - \delta, \hat{p}_M + \delta]) \geq c$.

**Proof**: Let $\omega$ be a path through the MDP $M$ and $D((x(t), y(t)), d(t))$ the corresponding uncertainty region as defined in Sec. VI. Let $q'(t)$ be any state trajectory as defined in Prop. 2. Also, let $\sigma^{\omega}$ and $\sigma^{q'}$ be the corresponding traces. Trace $\sigma^{\omega}$ can (i) satisfy $\phi$ and (ii) not satisfy $\phi$.

Let us first consider the former. If $\sigma^{\omega} = \phi$ from Prop. 3 it follows that $\sigma^{\omega} = \phi$. Under $\gamma$ the probability of generating $q'(t)$ is equivalent to generating path $\omega$ under $\mu_{det}$. Since under $\mu_{det}$ the probability that a path through the MDP $M$ generates a satisfying trace is $p_M$, it follows that the probability that the system given by Eqn. (1), under $\gamma$, will generate a satisfying state trajectory is also $p_M$.

To show that $p_M$ is the lower bound we need to consider the latter case. It is sufficient to observe that because of the conservative approximation of $D((x(t), y(t)), d(t))$ it is possible that $\sigma^{q'}$ satisfies $\phi$, even though $\sigma^{\omega}$ does not satisfy it. Therefore, it follows that the probability that system given by Eqn. (1), under the vehicle control strategy $\gamma$, generates a state trajectory that satisfies BLTL formula $\phi$, is bounded from below by $p_M$. The rest of the proof, i.e., $\Pr(p_M \in [\hat{p}_M - \delta, \hat{p}_M + \delta]) \geq c$, is given in [ZPC10].
The motion specification was:

- **Pick-up region**: to pick-up the load. After entering the **Pick-up region**, reach a **Test1 region** within 5 time units and stay in it at least 0.8 time units. Finally, after entering the **Test1 region** or the **Test2 region** reach a drop-off region within 4 time units to drop off the load. **Always avoid the unsafe regions**.

The specification translates to BLTL formula \( \phi \) in Eq. (10):

\[
\phi = -u_1 \leq 0.8 \land -u_2 \leq 0.8 \land \left( (u_1 \leq 1 \lor u_2 \leq 1) \land -u_3 \right).
\]

Two different environments are shown in Fig. 7. The estimated probability \( \hat{p}_e \) corresponding to environment A and B was 0.664 and 0.719, respectively. From Eq. (10) it followed that \( K = 9 \). The numerical values in the control synthesis procedure and the probability estimation procedure were as follows: \( N = 10000, h = 0.6, g = 0.6, \delta = 0.05, c = 0.95, \alpha = \beta = 1, \text{ and } e = 0.05 \). For both environments, we found the vehicle control strategy through the method described in Sec. VIII.

---

**E. Complexity**

As stated above, the size of the MDP \( M \) is bounded above by \((|U_t| \times |U_l| \times n_t \times n_l)^k\). Obviously, it can be expensive (in sense of memory usage) to store the whole MDP. Since our approach is sample-based, it is not necessary for the MDP to be constructed explicitly. Instead, a state of the MDP is stored only if it is sampled during the control synthesis procedure. As a result, during the execution, the number of states stored in the memory is bounded above by \( N \times K \times n \), where \( n \) is the number of iterations between the control synthesis and the probability estimation procedures.

The complexity analysis of the control synthesis part can be found in [HMZ+12] and the complexity analysis of BIE algorithm can be found in [ZPC10].

**IX. Case Study**

We considered the system given by Eqn. (1) and we used the numerical values corresponding to Dr. Robot’s x80Pro mobile robot equipped with two incremental encoders. The parameters were \( r = 0.085m \) and \( L = 0.295m \). To reduce the complexity, \( \{U_t \times U_l\} \) was limited to \( \{(1+L, 1-L), (1-L, 1-L), (1-L, 1+L)\} \), where the pairs of control inputs corresponded to a vehicle turning left at \( \frac{1}{2} \text{rad} \), going straight, and turning right at \( \frac{1}{2} \text{rad} \), respectively, when the forward speed is \( \frac{1}{4} \text{m/s} \).

**Measurement resolution**: To obtain the angular wheel velocity, the frequency counting method [PTPZ07] was used, i.e., the encoder pulses inside a given sampling period were counted. The number of pulses per revolution (i.e., the number of windows in the code track of the encoders) was 378 and the sampling period was set to \( \Delta t = 2.6s \). Thus, according to [PTPZ07] the measurement resolution was \( \Delta \epsilon = \frac{\Delta t}{2 \pi \times 378, 750} \approx 0.0064 \).

**Probability density functions**: We obtained the distributions through experimental trials. Specifically, we used control inputs from \( \{U_t \times U_l\} \) as the robot inputs and then measured the actual angular wheel velocities using the encoders. We obtained \( \epsilon_i^{\min} \) (\( \epsilon_i^{\max} \)) by taking the minimum (maximum) over \( \{\tilde{\epsilon}^1, \ldots, \tilde{\epsilon}^k\} \), \( \{\tilde{\epsilon}^{l+1}, \ldots, \tilde{\epsilon}^{l+k}\} \), \( j \in \{1, \ldots, k\} \), \( i \in \{r, l\} \), where \( \tilde{\epsilon}_i^j \), \( j \in \{1, \ldots, k\} \), \( i \in \{r, l\} \) was the noise interval, of length \( \Delta \epsilon_i \), determined from the \( j \)-th measurement of the encoder \( i \) and \( k \) was the total number of measurements. Note that \( n_i = \frac{\epsilon_i^{\max} - \epsilon_i^{\min}}{\Delta \epsilon_i}, i \in \{r, l\} \).

Finally, the probabilities for Eq. (4) that defined the probability density functions, were equal to the number of times a particular noise interval was measured over \( k \). For \( k = 150 \) (i.e., by using each control input from \( \{U_t \times U_l\} \) 50 times) we obtained \( -\epsilon_i^{\min} = \epsilon_i^{\max} = -\epsilon_i^{\min} = \epsilon_i^{\max} = 0.0096 \) and the corresponding probabilities.

The set of propositions was \( \Pi = \{\pi_u, \pi_p, \pi_{p1}, \pi_{p2}, \pi_d\} \) where \( \pi_u, \pi_p, \pi_{p1}, \pi_{p2}, \pi_d \) labeled the unsafe, pick-up, test1, test2 and the drop-off regions, respectively.

The motion specification was:

- **Start from an initial state** \( q_{init} \) and reach a **pick-up region** within 14 time units and stay in it at least 0.8 time units, to pick-up the load. After entering the **pick-up region**, reach a **test1 region** within 5 time units and stay in it at least 1 time units or reach a **test2 region** within 5 time units and stay in it at least 0.8 time units. Finally, after entering the **test1 region** or the **test2 region** reach a drop-off region within 4 time units to drop off the load. **Always avoid the unsafe regions**.

Since it is not possible to obtain the exact probability that the system given by Eqn. (1), under the vehicle control strategy, generates a satisfying state trajectory, in order to verify our result (Theorem 1), we performed multiple runs of BIE algorithm by simulating the system under the vehicle control strategy (using the same numerical values as stated above and by generating traces as described in Sec. IV).
We denote the resulting probability estimate as $\hat{p}_S$ and we compare it to $\hat{p}_M$.

### TABLE I

| Environment | $\hat{p}_M$ | $\hat{p}_S$ |
|-------------|-------------|-------------|
| Run 1       | Run 2       | Run 3       |
| A           | 0.664       | 0.847       | 0.832       | 0.826       |
| B           | 0.719       | 0.891       | 0.998       | 0.879       |

In Fig. 7 we show sample state trajectories and in Table I we compare the estimated probabilities obtained on the MDP, $\hat{p}_M$, with the estimated probabilities obtained by simulating the system, $\hat{p}_S$. The results support Theorem 1, since $\hat{p}_S$ is bounded from below by $\hat{p}_M$. The discrepancy in the probabilities is mostly due to the conservative approximation of the uncertainty region in Sec. VI. The Matlab code used to obtain the vehicle control strategy ran for approximately 2.2 hours on a computer with a 2.5GHz dual processor.

In Fig. 8 we show a sample run of the robot in environment A. A projector was used to display the environment and the state (position) trajectory was reconstructed using the OptiTrack system with eight cameras.

### X. DISCUSSION

We developed a feedback control strategy for a stochastic differential drive mobile robot such that the probability of satisfying a time constrained specification given in terms of a temporal logic statement is maximized. By mapping sensor measurements to a Markov Decision Process (MDP) we translate the problem to finding a control policy maximizing the probability of satisfying a Bounded Linear Temporal Logic (BLTL) formula on the MDP. The solution is based on Statistical Model Checking for MDPs and we show that the probability that the vehicle satisfies the specification is bounded from below by the probability of satisfying the specification on the MDP.

The key limitation of the proposed approach is the computation time. Since our algorithm is based on Statistical Model Checking for MDPs presented in [HMZ+12], to put the running time of our algorithm into perspective, we compare it to the running time of Statistical Model Checking for MDPs when dealing with the following motion planning study: each of the two robots living in a $20 \times 20$ grid world must pick up some object and then meet with the other robot within a certain time bound, while avoiding unsafe grids. At each point in time, either robot can try to move 10 grid units in any of the four cardinal directions, but each time a robot moves, it has some probability of ending up somewhere in a radius of 3 grid units of the intended destination. Statistical Model Checking for MDPs (when $c = 0.95$) solves this problem in approximately 20 minutes. Now, let us consider the algorithm and the case study presented in this paper. First, note that in order for Theorem 1 to hold, when generating a trace corresponding to an uncertainty region, we have to perform a computationally expensive step of taking the intersection between the uncertainty region and all of the regions of interest (see Def. 5, Sec. VII). Second, note that we are dealing with a system that is continuous both in space and time. Therefore, at each time step, when constructing an uncertainty region, the algorithm is required to perform multiple integrations of the system. Thus, the increase in the computational complexity, in order to have probabilistic guarantees for the original system, is the reason the running time of our algorithm (when $c = 0.95$) is approximately 2.2 hours.

Since sampling (i.e., generating traces) accounts for the majority of our runtime, future work includes improving the sampling performance and making the implementation fully parallel. Additionally, to address the problem of discrepancy between the probabilities obtained on the MDP and the probabilities obtained by simulating the system the future work also includes developing a less conservative uncertainty model.
XI. ACKNOWLEDGEMENTS

The authors gratefully acknowledge David Henrques from Carnegie Mellon University for comments on an earlier draft. Also, the authors would like to thank Benjamin Troxler, Michael Marrazzo and Matt Buckley from Boston University for their help with the experiments.

REFERENCES

[ABDS08] A. Abate, A. D’Innocenzo, Maria D. Di Benedetto, and S. Shankar Sastry. Markov Set-Chains as Abstractions of Stochastic Hybrid Systems, volume 1. Springer Verlag, 2008.

[BK08] C. Baier and J. P. Katoen. Principles of Model Checking. MIT Press, 2008.

[BM00] D. Balkcom and M. T. Mason. Time Optimal Trajectories for Bounded Velocity Differential Drive Robots. In IEEE International Conference on Robotics and Automation (ICRA) 2000, volume 3, pages 2499 – 2504, 2000.

[BMKV11] A. Bhatia, M. R. Maly, L. E. Kavraki, and M. Y. Vardi. Motion Planning with Complex Goals. Robotics Automation Magazine, IEEE, 18(3):55 –64, September 2011.

[CB12] I. Cizej and C. Belta. Probabilistically Safe Control of Noisy Dubins Vehicles. In International Conference on Intelligent Robots and Systems (IROS) 2012, October 2012.

[CGP99] E. Clarke, O. Grumberg, and D. A. Peled. Model Checking. The MIT Press, 1999.

[DABS08] A. D’Innocenzo, A. Abate, M. D. Di Benedetto, and S. Shankar Sastry. Approximate Abstractions of Discrete-Time Controlled Stochastic Hybrid Systems. In IEEE Conf. on Decision and Control, December 2008.

[DLB12] X. C. Ding, M. Lazar, and C. Belta. Receding Horizon Temporal Logic Control for Finite Deterministic Systems. In American Control Conference (ACC) 2012., June 2012.

[FMAQ98] T. Fraichard, R. Mermond, and R. Alpes-Gravir. Path Planning with Uncertainty for Car-Like Robots. In IEEE International Conference on Robotics and Automation (ICRA) 1998, pages 27–32, 1998.

[Gir07] A. Girard. Approximately Bisimilar Finite Abstractions of Stable Linear Systems. In International Conference on Hybrid Systems: Computation and Control (HSCC) 2007, pages 231–244, Berlin, Heidelberg, 2007. Springer-Verlag.

[HMZ+12] D. Henrques, J. Martins, P. Zuliani, A. Platzer, and E. M. Clarke. Statistical Model Checking for Markov Decision Processes. In 9th International Conference on Quantitative Evaluation of Systems (QEST) 2012, September 2012.

[JCL+09] S. Kumar Jha, E. M. Clarke, C. J. Langmead, A. Legay, A. Platzer, and P. Zuliani. A Bayesian Approach to Model Checking Biological Systems. In International Conference on Computational Methods in Systems Biology (CMSB) 2009, pages 218–234, 2009.

[JP09] A. A. Julius and G. J. Pappas. Approximations of Stochastic Hybrid Systems. IEEE Transactions on Automatic Control, 54(6):1193 –1203, 2009.

[KB08a] M. Kloetzer and C. Belta. Dealing with Non-Determinism in Symbolic Control. In Hybrid Systems: Computation and Control: 11th International Workshop, Lecture Notes in Computer Science, pages 287–300. Springer Berlin / Heidelberg, 2008.

[KB08b] M. Kloetzer and C. Belta. A Fully Automated Framework for Control of Linear Systems from Temporal Logic Specifications. IEEE Transactions on Automatic Control, 53(1):287 –297, 2008.

[KF08] S. Karaman and E. Frazzoli. Vehicle Routing Problem with Metric Temporal Logic Specifications. In IEEE Conference on Decision and Control (CDC) 2008., pages 3953 –3958, 2008.

[KGFP07] H. Kress-Gazit, G. E. Fainekos, and G. J. Pappas. Where’s Waldo? Sensor-Based Temporal Logic Motion Planning. In IEEE International Conference on Robotics and Automation (ICRA) 2007, pages 3116–3121, 2007.

[LAB12] M. Lahijanian, S. B. Andersson, and C. Belta. Temporal Logic Motion Planning and Control With Probabilistic Satisfaction Guarantees. IEEE Transactions on Robotics, 28(2):396 –409, 2012.

[LaV06] S. M. LaValle. Planning Algorithms. Cambridge University Press, Cambridge, U.K., 2006.

[LK04] S. G. Loizou and K. J. Kyriakopoulos. Automatic Synthesis of Multi-Agent Motion Tasks Based on LTL specifications. In Decision and Control, 2004. CDC. 43rd IEEE Conference on, volume 1, pages 153 – 158 Vol.1, Dec, 2004.

[PTFZ07] R. Petrella, M. Tursini, L. Peretti, and M. Zigliotto. Speed Measurement Algorithms for Low-Resolution Incremental Encoder Equipped Drives: A Comparative Analysis. In International Aegean Conference on Electrical Machines and Power Electronics (ACEMP) 2007, pages 780 –787, 2007.

[TTP06] P. Tabuada and G. J. Pappas. Linear Time Logic Control of Discrete-Time Linear Systems. IEEE Transactions on Automatic Control, 51(12):1862 –1877, 2006.

[WMT09] T. Wongpiromsarn, U. Topcu, and R. M. Murray. Receding Horizon Temporal Logic Planning for Dynamical Systems. In 48th IEEE Conference on Decision and Control (CDC) 2009, pages 5997 –6004, 2009.

[YTC+12] B. Yordanov, J. Tumova, I. Cerna, J. Barnat, and C. Belta. Temporal Logic Control of Discrete-Time Piecewise Affine Systems. IEEE Transactions on Automatic Control, 57(6):1491 –1504, 2012.

[ZPC10] P. Zuliani, A. Platzer, and E. M. Clarke. Bayesian Statistical Model Checking with Application to Simulink/Stateflow Verification. In International Conference on Hybrid systems: Computation and Control (HSCC) 2010, pages 243–252, 2010.